

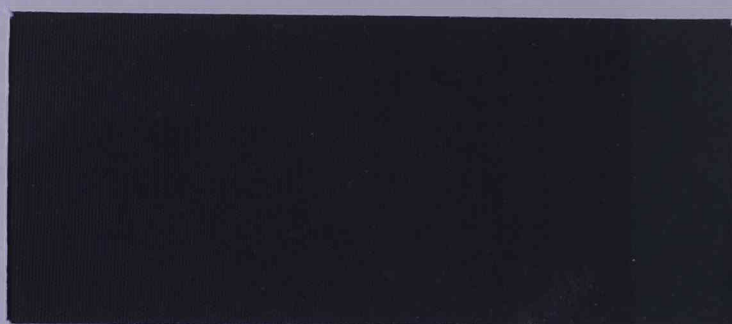


WORKING PAPERS

W.P. 85

COSTRUZIONE DELLE MATRICI DI FLUSSO
DELLE FORZE DI LAVORO PER LA
REGIONE PIEMONTE
STUDIO DI FATTIBILITA'

Luigi Varbella





W.P. 85

**COSTRUZIONE DELLE MATRICI DI FLUSSO
DELLE FORZE DI LAVORO PER LA
REGIONE PIEMONTE**

STUDIO DI FATTIBILITA'

Luigi Varbella

Aprile 1988



Si ringraziano il dr. Sanetti e il dr. Cascioli della Sezione Forze di Lavoro dell'ISTAT, il dr. Caperdoni del Servizio Statistica della Regione Lombardia, il dr. Miceli dell'IRES, il prof. Ricolfi del Dipartimento Scienze Sociali dell'Università di Torino e il dr. Viano dell'Osservatorio Regionale sul Mercato del Lavoro della regione Piemonte per i suggerimenti forniti e la disponibilità dimostrata.

INDICE

Capitolo 0.	INTRODUZIONE	pag.	1
Capitolo 1.	LE ESIGENZE E LE MOTIVAZIONI	"	7
	1.1. Tecnologia, complessità e accelerazio ne: instabilità	"	10
	1.1.1. Trasformazioni nell'assetto del sistema produttivo	"	19
	1.1.2. Modificazioni sociali e terri- toriali	"	23
	1.1.3. Transizione alla società post- industriale	"	26
	1.2. Nuovi strumenti d'indagine: ridefini- zione e riorientamento	"	32
Capitolo 2.	LE METODOLOGIE E LE ESPERIENZE	"	39
	2.1. I flussi di forza lavoro	"	41
	2.1.1. Risultati a scala nazionale	"	45
	2.1.2. La scala regionale	"	49
	2.2. I flussi da elaborazione di stock	"	55
	2.3. Le sperimentazioni su dati amministra- tivi in Piemonte	"	59
	2.3.1. La demografia industriale	"	61
	2.3.2. I flussi occupazionali e loca- lizzativi della industria mani- fatturiera	"	62
Capitolo 3.	CONCLUSIONI: I PROBLEMI E LA FATTIBILITA'	"	65
Bibliografia		"	67
Appendice 1	La procedura Istat	"	73
Appendice 2	Flussi da elaborazione di stock	"	87

0. INTRODUZIONE

In estrema sintesi, si può affermare che ogni approccio di indagine che riguardi gli assetti o le dinamiche del sistema economico, della società, del loro integrarsi o della loro dislocazione sul territorio conduce, a partire dall'inizio degli anni '80, invariabilmente a riconoscere, innanzitutto, che le cose si sono apprezzabilmente complicate.

Per di più la complicazione non riguarda tanto l'entità, pure cospicua, dei cambiamenti avvenuti o delle trasformazioni in atto, bensì la direzione o la qualità del mutamento che, in molte occasioni, presentano caratteristiche tali da svuotare di senso opinioni consolidate, sistemi di classificazione ed ordinamento dei fenomeni, relazioni di dipendenza fra le variabili e ciò sia al livello elementare di modo di formazione del senso comune, sia rispetto ad elaborati e sofisticati modelli di interpretazione e/o previsione dei fenomeni socio-economici.

In altri termini la vanificazione o la riduzione della capacità di interpretazione o spiegazione del nuovo, che si avverte utilizzando i tradizionali strumenti di indagine entro schemi consolidati, riguarda non solo e non tanto l'esigenza di descrivere le trasformazioni avvenute individuandone gli elementi determinanti ed essenziali, e le valenze di ciascuno di questi in rapporto al ruolo contingente oppure permanente, ma anche, e soprattutto, la possibilità di utilizzare l'analisi delle trasformazioni come fondamento per l'estrapolazione di tendenze future o la costruzione di compatibili scenari evolutivi.

Può apparire drastico questo modo di porre la questione, e certo i concetti sono stati estremizzati, ma unicamente in funzione e nell'intento di meglio isolarli e focalizzarli; non si tratta comunque di esagerazioni. A questo proposito si possono fare alcuni esempi.

Non è di poco conto il fatto che oggi si possa affermare che lavoro e non lavoro non sono più partizioni e che in base a tale schema non è più possibile classificare univocamente la popolazione

per posizione professionale data la maggiore e frequente presenza di soggetti definibili soltanto in rapporto a due o più posizioni contemporaneamente (studente e lavoratore, studente e disoccupato, disoccupato con o senza posto di lavoro, lavoratore precario in cerca di occupazione stabile, occupato in un settore di attività in cerca di occupazione in altro settore, ecc.): questa crescente duplicità o molteplicità di condizione non è neppure risolvibile in termini di prevalenza dell'una sulle altre perché si tratta di una rappresentazione dell'elevata mobilità tra diverse condizioni che interessa larghi strati di popolazione e che, al limite, appunto, si manifesta nello stesso soggetto proprio in termini di alternanza della prevalenza fra le condizioni coesistenti.

Un altro esempio può essere costituito dal riconoscere che oggi, nell'organizzazione dell'apparato produttivo industriale, gli effetti delle economie di scala non si manifestano tanto e soltanto in relazione al crescere delle dimensioni delle imprese, e della progressiva concentrazione al loro interno della maggior parte delle lavorazioni e delle fasi del ciclo produttivo, quanto piuttosto, e più direttamente, in rapporto al crescere del grado di integrazione fra più imprese, anche di dimensioni modeste, caratterizzate da un alto livello di specializzazione in singole lavorazioni o fasi rispetto alle quali l'organizzazione della produzione si è andata sempre più modularmente scomponendo.

Una ulteriore efficace esemplificazione può essere evidenziata citando l'ormai unanime riconoscimento che si è fortemente indebolita la relazione di diretta dipendenza dei livelli e/o dei tassi di sviluppo occupazionale rispetto ai livelli e/o ai tassi di sviluppo della produzione industriale e ciò non soltanto con riferimento al periodo più o meno lungo di riorganizzazione/ristrutturazione dell'apparato industriale in termini di effetto contingente e temporaneo, bensì come dato persistente anche dopo la compiuta attuazione dei programmi di riassetto produttivo e come caratteristica permanente della nuova organizzazione del sistema produttivo industriale.

Allora la complicazione, di cui si è fatto cenno all'inizio,

deriva dalla natura degli elementi che hanno determinato il cambiamento e che mostrano la capacità di incidere sui fondamenti stessi dell'assetto socio-economico complessivo.

L'elemento centrale è riconoscibile nella riconversione tecnologica, intesa come processo di sostituzione, a partire dal settore industriale, della tecnologia meccanica con la tecnologia elettronica. Ciò produce maggiore complessità e una accelerazione dei processi evolutivi, ovvero ad un maggior grado di flessibilità dell'organizzazione produttiva corrisponde una maggior articolazione sociale non solo in termini di ruolo, ma anche in senso spaziale, di localizzazione delle attività e delle residenze, ed il sistema socio-economico complessivo che ne risulta appare caratterizzato da un crescente grado di instabilità.

Questa instabilità deriva in primo luogo dalle mutate condizioni del lavoro e, in particolare, dalla rapida transizione delle condizioni di garanzia del posto di lavoro dal lungo al medio/breve periodo, anche nella componente di occupazione stabile con contratto a tempo indeterminato. Infatti, fino a tutti gli anni '70 si considerano come posti di lavoro garantiti non solo quelli nel settore pubblico, più direttamente e tradizionalmente caratterizzati in questo senso, ma anche, e a pari livello di garanzia di fatto, quelli nelle grandi imprese industriali. Un effetto non secondario delle trasformazioni della prima metà degli anni '80 è stato quello di ridimensionare vistosamente, non solo sul piano della percezione, le caratteristiche garantiste e di stabilità del lavoro nella grande impresa industriale, sia privata sia pubblica, e inoltre di mettere all'ordine del giorno la revisione non solo della condizione di immobilità dei posti di lavoro pubblici, ma anche della loro stessa stabilità, intesa come garanzia di continuità.

Un ulteriore indicatore dell'aumento del grado di instabilità del sistema produttivo è riconoscibile nella forte dinamica espansiva del numero di unità locali dell'industria manifatturiera, già evidenziato dal censimento 1981 rispetto al 1971 (+26,7% in Piemonte nel decennio), e tanto più rilevante in quanto si manifesta in corrispondenza ad un sensibile ridimensionamento del carico

occupazionale complessivo di riferimento (-4,1% gli addetti in Piemonte nel decennio 1971-1981). Questo risultato può essere ricondotto, sul piano delle determinanti, al processo di riconversione tecnologica, e il sottolineare l'effetto di accresciuta instabilità del sistema rappresenta il contrappunto problematico alle positive valutazioni in termini di maggiore vivacità imprenditoriale e di dispiegamento delle grandi e flessibili potenzialità produttive delle nuove tecnologie.

A fronte di tutto ciò si sono andati avvertendo sempre più chiaramente i limiti delle analisi basate su informazioni e dati raccolti in funzione della determinazione e quantificazione, ad una certa data, degli stock delle variabili prese in esame. E inoltre, se ci si limita a lavorare sui dati di stock, si tende ad articolare in maggior misura la disaggregazione dell'informazione e dell'interpretazione per cogliere la maggior complessità dei fenomeni e/o a ravvicinare le osservazioni sui fenomeni rilevandone l'entità con maggiore frequenza, e ciò per cogliere o inseguire l'accelerazione delle trasformazioni in atto.

Lungo questa via si può rispondere solo in parte alle nuove esigenze conoscitive che complessità e accelerazione nel divenire dei fenomeni socio-economici pongono: ovvero appare limitativo, oltre che faticoso, contrapporre alla maggiore instabilità strutturale una congiunturalizzazione dell'analisi. La soluzione più appropriata appare invece quella di qualificare l'analisi, arricchendola di valutazioni coerenti con il tipo di problemi che si presentano e che consistono, in sintesi, nell'individuazione ed interpretazione delle direzioni del cambiamento (le tendenze quindi, oltre che le entità) in assenza di un trend complessivo di indirizzo delle trasformazioni.

E' qui, e in questo modo, che si pone l'esigenza di misurare ed analizzare la realtà socio-economica anche in termini di flussi, per riuscire a cogliere le implicazioni qualitative che sottendono alle variazioni quantitative delle grandezze considerate. E' una questione che si pone al crocevia fra le riflessioni sugli effetti dei cambiamenti e la consapevolezza che occorre, di conseguenza, revisionare gli strumenti e le metodologie di analisi.

Questo tema viene qui affrontato, in particolare, in relazione all'esigenza avvertita di disporre di informazioni sui flussi anche a scala territoriale regionale, o comunque più articolata rispetto al livello nazionale, che costituisce l'unico ambito di elaborazione dei dati oggi sufficientemente dettagliato ed esauriente.

Ci si propone di individuare modi e forme che consentano di definire, scegliere ed adottare le opportune iniziative e le appropriate procedure atte alla realizzazione di strumenti di indagine siffatti attraverso un esame: in primo luogo, delle esigenze e motivazioni cui potrebbero rispondere informazioni organizzate e sistematiche sui flussi, a partire da quelli relativi alle forze di lavoro; in secondo luogo, delle metodologie disponibili per la loro realizzazione a scala territoriale regionale, tenendo conto anche di quelle definite per esaminare ambiti o problemi più circoscritti o particolari ed adottate in esperienze di ricerca orientate, in questi anni, ad approfondire elementi di conoscenza attraverso una analisi dei flussi; infine, dei problemi e delle difficoltà che l'applicazione di tali metodologie comporta, e dei limiti delle parziali esperienze esaminate, in modo da valutare la fattibilità, a quali condizioni e con quali forze, della matrice dei flussi della forza lavoro a livello regionale piemontese o, più in generale, di strumenti conoscitivi basati sulla realizzazione di archivi orientati in termini di flussi oltre che di stock.

1. LE ESIGENZE E LE MOTIVAZIONI

Un primo dato del problema, da cui si può partire, è quello della diffusa sensibilità che si è andata a coagulare intorno all'esigenza di disporre di informazioni, in particolare sul mercato del lavoro, organizzate in modo tale da consentire una analisi dei fenomeni, in particolare la mobilità di condizione rispetto al lavoro e la disoccupazione, in termini di flussi.

Le testimonianze in merito sono molteplici e provengono da ricercatori ed osservatori che operano in diverse sedi, accademiche, istituti di ricerca, uffici amministrativi, eppure convergono nel sottolineare l'opportunità di valorizzare tale strumento conoscitivo. Si possono citare Carmignani (1986) "abbiamo cercato, come Cespe di fare dei flussi un'opzione metodologica più che un indicatore congiunturale" e poi Pugliese (1986) che, pur in polemica con Carmignani ed Accornero nel merito della valutazione dei caratteri della disoccupazione in Italia, scrive: "va infatti riconosciuto ad Accornero ed al Cespe il merito di aver posto con forza l'attenzione ai flussi per l'analisi del mercato del lavoro".

Fin qui si tratta di testimonianze annotate a margine di studi e valutazioni su fenomeni esaminati a livello nazionale, e a quel livello i dati sono disponibili: sono quelli sui flussi di forza lavoro ormai pubblicati annualmente dall'Istat a partire dal 1981. Non solo, ma sono state altresì sperimentate efficaci metodologie di utilizzo di tali dati in funzione della definizione del grado di mobilità intersettoriale delle forze di lavoro (MAROTTA e PUGLIESE, 1986).

Le stesse esigenze, con pari sensibilità e forza, vengono poste anche da chi si occupa di problemi del mercato del lavoro o, più in generale, di dinamiche socio-economiche a livello regionale, pur se l'ambito territorialmente più circoscritto di indagine costituisce di per sé una riduzione del grado di complessità derivante dai dualismi e dalle specificità locali che confluiscono nel dato definito a livello nazionale complessivo. Si può qui citare l'affermazione, da parte

dell'Osservatorio sul mercato del lavoro della Regione Piemonte, di una "necessità di accompagnare le analisi in termini di stock, con analisi in termini di flusso, evidenziando i cambiamenti intervenuti in una certa popolazione in un determinato arco di tempo" (MAUGERI, 1987) nel contesto di una indagine sulle caratteristiche dell'offerta di lavoro nell'Area Metropolitana Torinese.

Ma a scala regionale le informazioni sui flussi di forza lavoro non sono disponibili, e allora si vanno moltiplicando i tentativi di costruire il dato statistico utilizzando come fonti alcuni archivi rilevati per finalità primarie di tipo amministrativo. Tali tentativi sono sempre finalizzati a riempire quel vuoto di informazione sui movimenti e sulle tendenze che determinano la transizione nel tempo da una composizione strutturale di stock ad un'altra, sensibilmente differenziata; e sono motivati dall'esigenza, costantemente espressa nella presentazione di tali lavori, di approfondire i problemi che un semplice confronto fra le attività definite al tempo iniziale e finale contribuisce più a porre ed evidenziare, piuttosto che a spiegare.

A questo proposito si possono citare, per esemplificare, due studi sulla dinamica dell'apparato produttivo piemontese che ricavano i dati di base rispettivamente dagli archivi INPS e dal Registro ditte delle CCIAA confrontati con gli archivi UPLMO e li elaborano in termine di flussi. Si tratta, nel primo caso, della ricerca sulla demografia industriale del Piemonte (CONTINI, ASCOLI, REVELLI, 1986) che analizza il recente forte dinamismo delle unità locali industriali, in particolare definendone i tassi di natalità e mortalità e costruendo le matrici di transizione per classi di addetti; nel secondo caso, dello studio sui flussi occupazionali e localizzativi dell'industria manifatturiera in provincia di Torino nel periodo 1980-1985 (IRES del Piemonte, 1987) che si propone di scomporre e quantificare, mantenendo un puntuale e dettagliato riferimento di localizzazione sul territorio, le diverse componenti, sia di crescita che di flessione, all'interno delle divergenti dinamiche generali delle unità locali e degli addetti.

Nell'unico caso di realtà regionale, la Lombardia, ove è già stata avviata una sperimentazione delle metodologie di stima dei movimenti

della popolazione da una condizione all'altra all'interno del mercato del lavoro (con la realizzazione delle matrici di flusso regionali 1980-1983), la disponibilità dei dati di flusso ha immediatamente dimostrato la sua ampia potenzialità interpretativa (ROSTI, 1986) al fine di cogliere le trasformazioni strutturali dell'occupazione che pure si manifestano, con effetti rilevanti, pur senza modificare la stabilità complessiva dei livelli, definiti in base al tasso di occupazione.

Ed è proprio in un contesto di questo tipo che i soli dati di stock risultano insufficienti ed inefficaci per interpretare la dinamica del mercato del lavoro perché assumono maggiore rilevanza quegli aspetti di movimento intersettoriale degli occupati, di transizione bidirezionale tra occupazione e disoccupazione e tra forze e non forze di lavoro che sono proprio quelli meglio evidenziati da dati elaborati in termini di flussi.

Ciò può contribuire, d'altra parte, ad individuare il ritardo nella sperimentazione o meglio il diverso approccio alla metodologia dei flussi, su problemi specifici e non in generale, del Piemonte rispetto alla Lombardia. In Piemonte, infatti, il forte peso dell'industria e, al suo interno, del settore dell'auto, e della grande dimensione d'impresa, ha fatto sì che gli effetti di crisi occupazionale, prodotti dalle trasformazioni della prima metà degli anni '80, fossero chiaramente percepibili anche al livello di stock, mentre l'esigenza di valutare il cambiamento in termini di flussi si pone piuttosto in relazione al problema delle modificazioni nella struttura dell'apparato industriale manifatturiero e della sua ridislocazione territoriale e dimensionale.

Le motivazioni di queste nuove esigenze di strumenti più sofisticati e flessibili di analisi delle dinamiche del mercato del lavoro hanno però carattere generale, nel senso che sono tutte riconducibili al maggior grado di instabilità della realtà socio-economica complessiva negli anni '80, instabilità intesa come sintesi del cambiamento avvenuto, a partire dalle trasformazioni tecnologiche che introducono maggiore flessibilità e potenzialità produttive e relazionali, ma anche, nel contempo, maggiore complessità

ed una sensibile accelerazione dei processi evolutivi.

Di qui la considerazione di come sia opportuno dotarsi di una strumentazione statistica in termini di flussi non occasionale o costruita volta a volta in relazione ad approfondimenti di singoli problemi, ma avente carattere generale e permanente. Appare allora adeguata la proposta di costruzione delle matrici di flusso ricavabili dalle rilevazioni trimestrali delle forze di lavoro, in quanto consentono di ricostruire un vero e proprio bilancio dinamico della popolazione disaggregata in base alle diverse condizioni professionali. E di ciò si avvertono motivate esigenze di disponibilità a scala regionale.

1.1 Tecnologia, complessità e accelerazione: instabilità

Può sorgere una perplessità che appare importante evocare, e cercare di confutare, perché può togliere forza alle motivazioni della proposta di adozione di una metodologia di analisi socio-economica basata su dati di flusso. Si tratta della tendenza a considerare questo tipo di approccio tutt'al più come un utile approfondimento dei problemi, capace di mettere meglio a fuoco alcune dinamiche "micro", di cogliere dettagli e curiosità, ma sostanzialmente con funzioni di semplice arricchimento ed integrazione rispetto alle tradizionali analisi di struttura basate su dati di stock.

Risulterebbe allora arbitrario il far derivare le esigenze e le motivazioni dell'analisi di flusso dalla natura delle cause delle trasformazioni della prima metà degli anni '80, e improponibile una individuazione di tale strumento in termini di più efficace chiave interpretativa al fine di comprendere e valutare la portata degli effetti di tali trasformazioni.

La tesi che qui si intende viceversa sostenere è che la questione del rapporto fra analisi degli stock e analisi dei flussi si colloca, oggi, al crocevia, ad un punto di svolta in cui si intersecano le

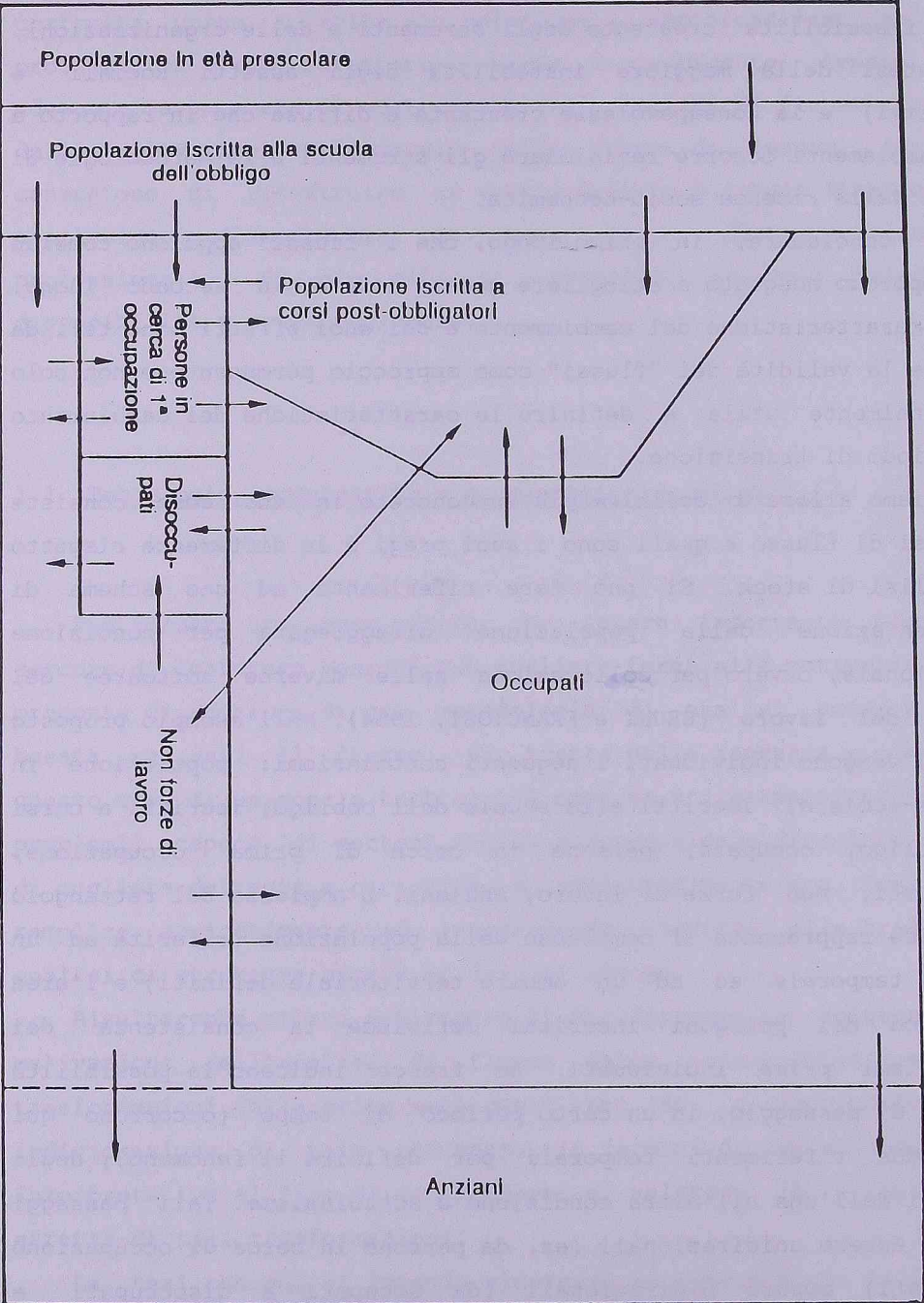
riflessioni sugli effetti dei cambiamenti profondi, tecnologici, sociali, economici, di organizzazione delle attività produttive e di servizio (sotto il segno della tendenza alla complessità dei fenomeni, della flessibilità crescente degli strumenti e delle organizzazioni, in sintesi della maggiore instabilità degli assetti sociali e produttivi) e la consapevolezza crescente e diffusa che in rapporto a tali cambiamenti occorre revisionare gli strumenti e le metodologie di analisi della ricerca socio-economica.

Per concludere, in primo luogo, che i "flussi" appaiono come un primo approdo adeguato a sciogliere questo nodo e, in secondo luogo, che le caratteristiche del cambiamento e dei suoi effetti sono tali da esaltare la validità dei "flussi" come approccio permanente e non solo occasionalmente utile a definire le caratteristiche del cambiamento nel periodo di transizione.

Vediamo allora di definire più in concreto in che cosa consiste l'analisi di flusso e quali sono i suoi pregi e le differenze rispetto all'analisi di stock. Si può fare riferimento ad uno schema di rappresentazione della popolazione disaggregata per condizione professionale, ovvero per collocazione nelle diverse sottoaree del mercato del lavoro (BRUNI e FRANCIOSI, 1984). Nell'esempio proposto (fig. 1) vengono individuati i seguenti sottoinsiemi: popolazione in età pre-scolare, iscritti alla scuola dell'obbligo, iscritti a corsi post-obbligo, occupati, persone in cerca di prima occupazione, disoccupati, non forze di lavoro, anziani. L'ampiezza del rettangolo di cornice rappresenta il complesso della popolazione (riferita ad un momento temporale ed ad un ambito territoriale definiti) e l'area geometrica dei poligoni inscritti definisce la consistenza dei sottoinsiemi prima individuati. Le frecce indicano le possibilità teoriche di passaggio, in un certo periodo di tempo (occorrono qui almeno due riferimenti temporali per definire il fenomeno), degli individui dall'una all'altra condizione o sottoinsieme. Tali passaggi possono essere unidirezionali (es. da persone in cerca di occupazione a occupati) oppure bidirezionali (da occupati a disoccupati e viceversa).

In tale schema le aree individuano gli stock e la loro natura di

Figura 1 - Schema di dislocazione della popolazione rispetto al mercato del lavoro



tratto da: M. Bruni e F.B. Franciosi, Il mercato del lavoro in Italia: una analisi di flusso, in: L'offerta di lavoro in Italia, a cura di M. Schenkel, Marsilio editori, 1984.

variabile intrinsecamente statica, ovvero assimilabile ad una rappresentazione fotografica della popolazione; le frecce individuano i flussi e la loro natura di variabile intrinsecamente dinamica, ovvero assimilabile ad una rappresentazione cinematografica della popolazione (1). Un'altra immagine capace di sintetizzare il rapporto stock-flussi può derivare dall'associazione con il concetto di assoluto-relativo, nel senso che mentre gli stock definiscono l'entità strutturale della realtà osservata, i flussi rappresentano le tendenze dinamiche della sua modificazione. Il punto importante che si ricava dall'associazione flussi-relativo consiste nella determinazione dell'ambito di validità del dato, nel senso che i flussi, per essere utili a spiegare l'evoluzione del fenomeno osservato, debbono sempre essere riferiti ad uno stock di partenza; diversamente si corre il rischio, comune ogni qualvolta si utilizzano esclusivamente dati relativi, di perdere di vista il rapporto con il valore assoluto, ovvero con la realtà.

Detto questo, si può osservare che l'analisi di stock consente un tipo di approccio al problema in grado soltanto di definire la consistenza della popolazione e dei suoi sottoinsiemi ad un dato tempo e, disponendo di due rilevazioni successive, le variazioni di tali consistenze in un intervallo di tempo. L'analisi di flusso consente invece di descrivere dettagliatamente tali variazioni e, con ciò, di spiegarle piuttosto che determinarne l'entità, anche se questo dato è egualmente ricavabile in quanto le variazioni di stock in un intervallo di tempo possono essere definite come saldo tra i flussi in uscita e i flussi in entrata rispetto alla stessa popolazione o sottoinsieme di essa.

(1) Secondo la lettura dello schema proposta da Bruni e Franciosi (1984): "Il mercato del lavoro può essere osservato facendo uso di due diverse prospettive: la prima di tipo statico, la seconda dinamica per natura e che sta alla prima come la ripresa cinematografica sta alla fotografia".

Per conseguenza, il disporre dei dati di flusso consente di trasferire e, nel contempo, meglio indirizzare lo sforzo interpretativo della valutazione delle variazioni di stock all'analisi dei movimenti fra i vari sottoinsiemi che tali variazioni contemporaneamente spiegano e determinano. L'arricchimento dell'informazione deriva dalla scomposizione del dato di variazione temporale che i "flussi" consentono definendo non solo il saldo (flusso netto assimilabile alla variazione di stock nel periodo corrispondente), ma anche le variazioni in entrata e in uscita da ciascun sottoinsieme (flussi lordi). Un esempio di tale arricchimento può essere ricavato prendendo in esame i dati relativi al flusso lordo delle entrate nell'area degli occupati (domanda di flusso occupazionale) suddivise per settore di attività rilevate, nel periodo 1981-1986 a livello nazionale.

TABELLA 1 (1)

SETTORI DI ATTIVITA'	STRUTTURA DELLE ENTRATE (DOMANDA DI FLUSSO) PER SETTORE DI ATTIVITA' NEL PERIODO 1981-1986 IN ITALIA	
	VALORI ASSOLUTI	VALORI PERCENTUALI
Agricoltura	116	3.7
Industria	620	19.7
Altre attività	2412	76.6
TOTALE	3148	100.0
MOVIMENTI INTERNI	379	12.0
TOTALE ENTRATE	2769	88.0

(1) Dati tratti dalla relazione presentata da M.Bruni al seminario "Mercato del lavoro: realtà statica o dinamica?", Unione Industriale Torino, Torino 6 ottobre 1987.

Il dato riferito alle altre attività in tabella 1 (76,6%) ci dice che nell'ultimo quadriennio 3 nuovi posti di lavoro su 4 vengono attivati dal settore terziario e si può interpretare come la variazione marginale che trascina il dato medio di stock dal 49,5% del 1981 al 56,6% del 1986 (cfr. tabella 2).

TABELLA 2

SETTORI DI ATTIVITA'	STRUTTURA SETTORIALE DELL'OCCUPAZIONE IN ITALIA			
	1981		1981	
	VALORI ASSOLUTI	VALORI PERCENTUALI	VALORI ASSOLUTI	VALORI PERCENTUALI
Agricoltura	2.732	13.3	2.241	10.7
Industria	7.647	37.2	6.821	32.7
Altre attività	10.165	49.5	11.794	56.6
TOTALE	20.544	100.0	20.856	100.0

Fonte ISTAT, rilevazioni trimestrali delle forze di lavoro

L'analisi dei flussi sembra dunque costituisce un valido strumento per prevedere non certo quale sarà il futuro andamento dell'occupazione, bensì quelle che possono essere le conseguenze di un aumento o di una diminuzione dello stock occupazionale definiti in opportuni scenari di riferimento, oltre che spiegare, come l'esempio proposto evidenzia, le variazioni occupazionali che si sono manifestate quantificando le tendenze che le hanno determinate.

Quando le trasformazioni sono di rilevante entità ed alcuni processi subiscono una accelerazione molto elevata, come è avvenuto nella prima metà degli anni '80 in Italia ma in particolare in Piemonte, cresce l'interesse a conoscere le variazioni marginali più che non quelle che si vanno consolidando nel medio periodo, e tale esigenza deriva da un'altra caratteristica essenziale dei "flussi" che individuano tanti più passaggi bidirezionali rispetto all'insieme

considerato quanto più è breve l'intervallo di tempo di riferimento. E ciò in quanto nel medio-lungo periodo sfugge alla rilevazione tutta una serie di passaggi temporanei, non per questo meno significativi, che certamente rappresentano una entità consistente, come evidenziano i dati riportati in tabella 3. Si tratta degli stessi cambiamenti di condizione (passaggi) riferiti allo stesso periodo, ma, una volta rilevati in rapporto all'intervallo di tempo complessivo (anno) e, un'altra volta, rilevati in rapporto a ciascuno dei quattro trimestri e poi sommati.

TABELLA 3 - Cambiamenti di condizione (1)

Valori annuali (media 1979/80 e 1983/84) su base:		
	trimestrale	annua
UOMINI		
occupati	2.046	896
forza lavoro	2.051	894
DONNE		
occupati	2.246	891
forza lavoro	2.913	1.107
TOTALE		
occupati	4.310	1.787
forza lavoro	4.963	2.001

Appare chiaramente una consistente divergenza che comunque non è occasionale, ovvero non dipende dalle particolari situazioni e dinamiche cui si riferiscono i dati riportati in tabella 3. Si tratta di una importante e permanente caratteristica delle informazioni sui flussi, di cui occorre tener conto nel loro utilizzo. Ciò può essere meglio colto intuitivamente notando, sempre con riferimento ai dati esemplificativi proposti, che tutto il sistema di passaggi dalla

(1) Dati tratti dalla relazione presentata da M.Bruni al seminario "Mercato del lavoro: realtà statica o dinamica?", citata.

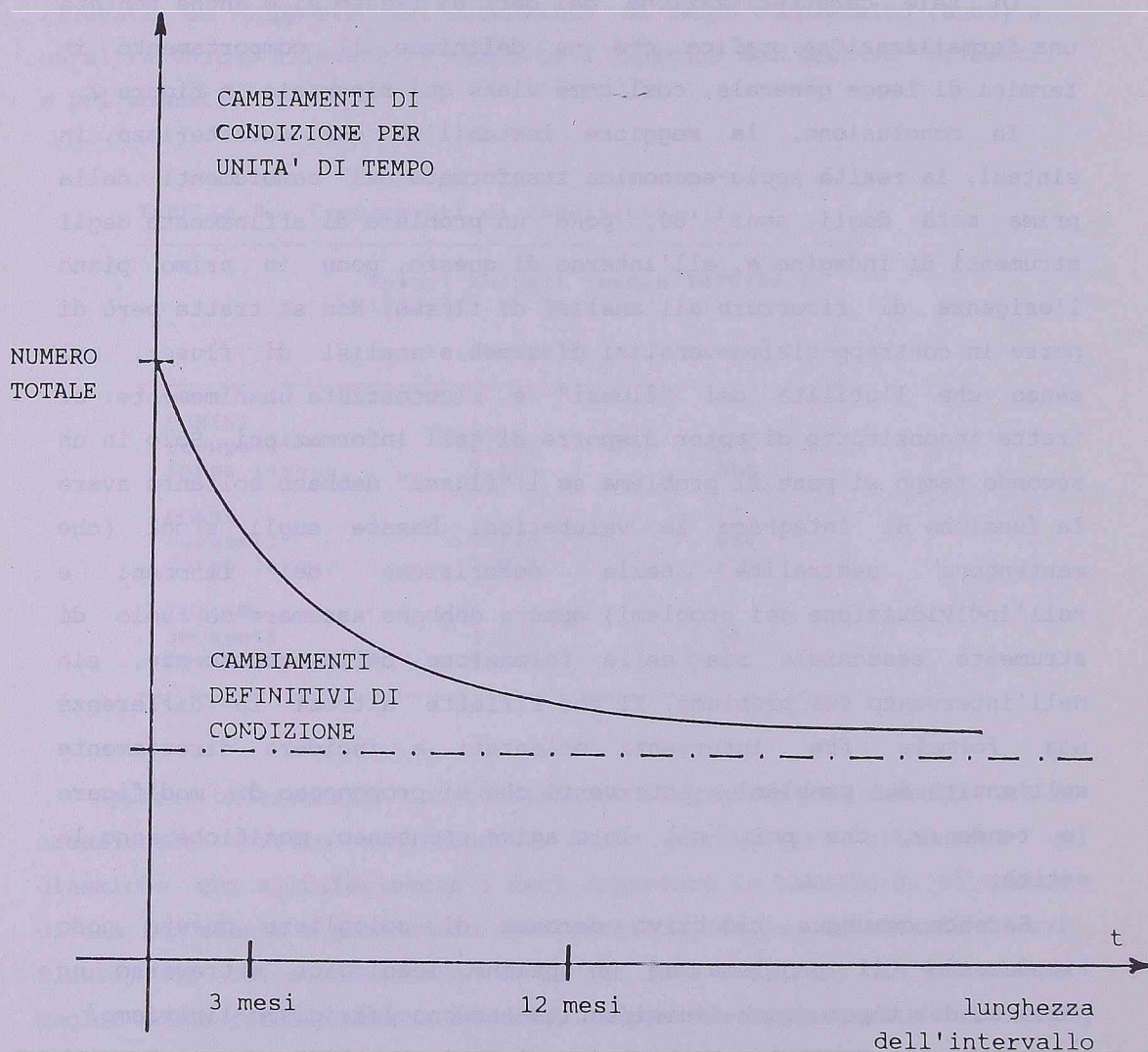
condizione di disoccupazione a quella di occupazione, e poi ancora a quella di disoccupazione, che caratterizza i lavoratori occupati solo stagionalmente e che a inizio e fine periodo (anno) risultano disoccupati, sfugge alla rilevazione quando questa viene effettuata su base annua, mentre viene evidenziata dalle indagini svolte su base trimestrale.

Di tale caratterizzazione dei dati di flusso si è anche tentata una formalizzazione grafica che ne definisce il comportamento in termini di legge generale, così come viene qui riportato in figura 2.

In conclusione, la maggiore instabilità che caratterizza, in sintesi, la realtà socio-economica trasformata dai cambiamenti della prima metà degli anni '80, pone un problema di affinamento degli strumenti di indagine e, all'interno di questo, pone in primo piano l'esigenza di ricorrere all'analisi di flusso. Non si tratta però di porre in contrapposizione analisi di stock e analisi di flusso, nel senso che l'utilità dei "flussi" è riconosciuta unanimemente: si tratta innanzitutto di poter disporre di tali informazioni. Solo in un secondo tempo si pone il problema se i "flussi" debbano soltanto avere la funzione di integrare le valutazioni basate sugli stock (che mantengono centralità nella descrizione dei fenomeni e nell'individuazione dei problemi) oppure debbano assumere un ruolo di strumento essenziale sia nella formazione delle conoscenze, sia nell'intervento sui problemi. Il che riflette altresì la differenza non formale fra interventi orientati a incidere direttamente sull'entità dei problemi e interventi che si propongono di modificare le tendenze, che poi, nel loro agire spontaneo, modificheranno le entità.

Sarebbe comunque riduttivo cercare di sciogliere questo nodo riducendo il problema ad un'opzione ideologica attraverso un percorso del tipo: stock-flussi/entità-tendenze/dirigismo-liberismo/stato/mercato. E ciò in quanto la maggiore instabilità e complessità dei fenomeni non deriva soltanto da una fase di prevalenza del ruolo del mercato, ma dipende dalle caratteristiche delle trasformazioni degli anni '80 che, per fare riferimento ancora allo schema riportato in figura 1, non incidono tanto sulla dimensione delle aree

Figura 2 - Flussi: numero di passaggi evidenziati in rapporto all'intervallo temporale prescelto.



tratto da: M. Bruni, relazione presentata al seminario "Mercato del lavoro: realtà statica o dinamica?", Unione Industriale Torino, Torino, 6 ottobre 1987.

sottoinsiemi, quanto piuttosto nel potenziare la portata delle "condutture" simboleggiate dalle frecce, ovvero nel determinare un aumento del numero di passaggi dall'una all'altra condizione nel sottolinearne la bidirezionalità effettiva e non solo potenziale.

Il fatto che questo maggior dinamismo si affermi sempre più come carattere permanente della realtà impone evidentemente non solo che quest'ultima venga studiata con strumenti adeguati che non possono prescindere da un'analisi dei flussi, ma anche che ogni tipo e modo di intervento sui problemi deve rapportarsi a tali caratteristiche.

Si può infine ancora brevemente esaminare come le esigenze e le motivazioni di disporre di dati sui flussi si manifestino e si evidenzino in relazione a ciascuno dei principali effetti del cambiamento e, in particolare rispetto alle trasformazioni nell'assetto del sistema produttivo; alle modificazioni sociali e territoriali; alla transizione alla società post-industriale.

1.1.1 Trasformazioni nell'assetto del sistema produttivo

La struttura dell'apparato produttivo industriale piemontese si è sensibilmente modificata, già a partire dagli anni '70, in direzione di una maggiore articolazione complessiva.

Le caratteristiche generali del cambiamento emergono chiaramente già dal confronto fra i dati dei censimenti industriali al 1971 e al 1981 e si possono sintetizzare in una generalizzata (settoriale e territoriale) espansione delle iniziative cui corrisponde una altrettanto generalizzata riduzione dell'attivazione occupazionale complessiva. La flessione degli aspetti è poi direttamente correlata con l'ampiezza occupazionale della dimensione aziendale: nelle unità più grandi (oltre 1.000 addetti) si perdono 63.000 posti di lavoro, mentre la piccola impresa (fino a 50 addetti) aumenta di circa 30.000 unità la propria forza lavoro; il saldo occupazionale complessivo dell'industria manifatturiera è pari a -32.000 addetti.

In estrema sintesi, meno addetti, occupati in più unità, più uniformemente distribuite sul territorio regionale.

L'insieme di questi effetti già pone, di per sè, un problema di parallela articolazione e qualificazione dell'analisi. E' sufficiente poi richiamare anche sommariamente le caratteristiche del processo di ristrutturazione, riorganizzazione e riconversione produttiva del settore industriale, per avere conferma che tali effetti e tale esigenza hanno carattere permanente e non occasionale.

Gli effetti enunciati si possono infatti ricondurre essenzialmente a due fenomeni:

- 1) la crisi occupazionale della grande dimensione d'impresa, dovuta alle crescenti difficoltà riconducibili al basso grado di flessibilità, sia operativa sia di mercato, in una fase di particolare accelerazione dei mutamenti economici, tecnologici e sociali;
- 2) la riqualificazione della funzione produttiva delle piccole e medie imprese che, a partire dalla crisi del 1975, hanno contribuito in modo decisivo a sostenere l'apparato produttivo piemontese sia in termini occupazionali sia, in particolare, in termini di competitività dei prodotti.

Il primo fenomeno ha provocato un ridimensionamento , in alcuni casi consistente anche dell'apparato produttivo, ma soprattutto dei livelli occupazionali, nei tradizionali poli di produzione industriale, caratterizzati da una particolare concentrazione di insediamenti di grandi aziende. In particolare la riduzione dei posti di lavoro nelle grandi imprese motrici regionali si manifesta come primo effetto della rapida diffusione della innovazione tecnologica, la cui origine può essere datata al 1980.

Sul versante dell'impresa minore, al contrario, le possibilità offerte dalle nuove tecnologie elettroniche, maggiormente compatibili con la scala dimensionale di tale tipo d'impresa, hanno creato le condizioni per lo sviluppo di nuova imprenditorialità.

Questa favorevole condizione di base si è manifestata in un particolare momento economico, caratterizzato dalla volubilità della domanda sui mercati sia dei prodotti finali di consumo, sia dei prodotti strumentali finali ed intermedi che costituiscono le

principali componenti dell'export non solo italiano, ma anche piemontese. L'elevato grado di flessibilità d'uso dei fattori produttivi, la capacità di rapido adattamento alle nuove e congiunturalmente mutevoli condizioni di mercato, lo stesso processo di decentramento produttivo innescato, a metà degli anni '70, per iniziativa della grande impresa, hanno costituito, per la piccola e media dimensione, occasioni di crescita, del numero delle unità produttive soprattutto e, principalmente per questa via, anche nei livelli occupazionali complessivi, a tassi di sviluppo molto elevati.

Ma il risultato più importante, che qualifica il nuovo ruolo esercitato nel sistema produttivo da questa tipologia d'impresa, è la dilatazione degli ambiti operativi, che si sono estesi da ristretti spazi interstiziali non occupati dalla grande dimensione, fino alle produzioni tecnologicamente avanzate.

Ciò apre alla piccola e media impresa una nuova prospettiva entro cui assume un ruolo non subordinato rispetto alle strategie delle imprese maggiori e non marginale rispetto ai settori e ai mercati. Si pone, anzi, come elemento centrale del nuovo assetto produttivo a specializzazione flessibile che si è andato affermando attraverso il processo di scomposizione modulare dell'apparato industriale.

Per approfondire questi temi non è direttamente utile il riferimento alle matrici di flusso delle forze di lavoro, ma occorre fondare l'analisi su dati relativi al sistema delle imprese, rispetto al quale, però, non vengono effettuate rilevazioni sistematiche a fini statistici, salvo quelle di tipo censuario. Le rilevazioni trimestrali delle forze lavoro sono infatti indagini sulle famiglie, rilevano gli occupati rispetto al luogo di residenza e non rispetto al luogo di lavoro (unità locali) e possono quindi fornire, in relazione al sistema delle imprese, solo indicazioni indirette ricavabili dalla dinamica del lavoro indipendente assunto come indicatore del grado di vivacità imprenditoriale.

Perciò i tentativi di meglio penetrare questi problemi orientando la ricerca alla determinazione dei "flussi" che sono stati messi in campo in Piemonte hanno più appropriatamente fatto riferimento agli archivi di dati amministrativi sulle imprese, come già si è accennato.

In concreto si è trattato di quantificare il fenomeno di nati-mortalità delle imprese e determinare le matrici di transizione per classi di dimensione occupazionale (CONTINI, ASCOLI, REVELLI, 1986, su dati INPS), oppure elaborare le informazioni in termini di flussi occupazionali e localizzativi delle unità locali industriali (IRES del Piemonte, 1987, su dati Camera di Commercio).

E' interessante notare come anche in questo caso si evidenzia l'utilità di disporre di informazioni sistematiche sulla dinamica delle unità locali in termini di flusso.

A tale interesse intende dare una risposta complessiva la nuova serie della pubblicazione MOVIMPRESE (movimento anagrafico delle imprese italiane) realizzata a cura del CERVED. Oggetto di tale pubblicazione (ora semestrale) è la periodica rilevazione dei dati di natalità e mortalità delle imprese a scala nazionale, regionale e provinciale e con riferimento ad un elevato livello di dettaglio di classificazione delle attività economiche (le 66 voci corrispondenti alle classi di attività definite nello schema adottato dall'ISTAT in occasione del censimento 1981). I dati di base sono ricavati dai registri delle Camere di Commercio. Il principale limite di tale iniziativa, se la si vuole utilizzare per studiare la struttura e l'evoluzione del sistema produttivo, è costituita dal fatto che "le aziende plurilocalizzate ossia, presenti in più province, sono state conteggiate una sola volta ed attribuite alla provincia nella quale l'iniziativa imprenditoriale ha la sede principale"; ragione per cui "l'attuale formulazione di MOVIMPRESE si rivolge direttamente all'analisi dinamica del fenomeno imprenditorialità e non alla quantificazione dell'apparato produttivo e di servizio locale" (1). Movimprese rappresenta comunque l'unico strumento informativo sistematico attualmente disponibile in questo ambito di interesse ed il più importante esempio di utilizzo di una fonte amministrativa al fine di predisporre serie statistiche di dati. La sua nuova serie rappresenta inoltre un momento di evoluzione nella predisposizione di informazioni in qualche modo organizzate in termini di flusso per consentirne una finalizzazione d'uso non più orientata soltanto all'analisi congiunturale ma anche a cogliere le nuove caratteristiche

(1) Cfr. CERVED; MOVIMPRESE, movimento anagrafico della impresa italiana; n. 1; 1985; pag. 135.

generali del cambiamento; e non a caso tale evoluzione si manifesta "a metà degli anni '80, i quali presentano più di una diversità rispetto al decennio precedente" (2).

1.1.2 Modificazioni sociali e territoriali

Se le modificazioni sono imponenti, come si è visto, nell'assetto del sistema produttivo industriale e, per diretto riflesso, all'interno del mercato del lavoro, le conseguenze sulla dislocazione sociale e territoriale della popolazione non sono di minore portata in questo scorcio di anni '80.

In Piemonte è particolarmente evidente la riduzione del peso e del ruolo della classe operaia nella popolazione e nella società, misurata innanzitutto dal forte ridimensionamento dell'occupazione nel settore industriale manifatturiero, diffuso in ogni provincia e particolarmente accentrato nell'area centrale (provincia di Torino) ove si ha una maggiore concentrazione degli stabilimenti di grandi dimensioni:

TABELLA 4 - Addetti nell'industria manifatturiera piemontese a scala provinciale 1980-1985 (Istat, rilevazioni trimestrali delle forze di lavoro)

PROVINCIA	ADDETTI (000)		$\Delta\%$	QUOTE DI COMPOSIZIONE % SUL TOTALE OCCUPATI		
	1980	1985		% 1980	% 1985	Differenza
AL	55	45	-18.2	28.5	24.0	-4.5
AT	22	19	-13.6	22.9	22.2	-0.7
CN	66	60	-9.1	27.7	24.4	-3.3
NO	87	71	-18.4	42.0	34.8	-7.2
TO	457	349	-23.6	46.8	38.5	-8.3
VC	75	67	-10.7	44.9	40.0	-4.9
PIEMONTE	762	611	-19.8	40.6	34.1	-6.5

(2) Cfr. CERVED; MOVIMPRESE, cit., pag. 11.

La trasformazione non si manifesta soltanto in termini di riduzione del peso degli addetti manifatturieri pur se è così rilevante da prospettare scenari inimmaginabili ancora alla fine degli anni '70: "Forse non ci sarà più una questione operaia..... Per dirla in termini drastici,.....: la questione contadina sta scomparendo per la scomparsa dei contadini; in un futuro non lontano la questione operaia probabilmente verrà superata con la tendenziale scomparsa degli operai" (Sylos Labini, 1986). Si manifesta anche, e con pari intensità, all'interno dell'area - residua di occupazione manifatturiera: "nell'industria gli impiegati veri e propri rappresentavano il 10% dei lavoratori; oggi quella quota si è quasi triplicata. E se, nel passato, nelle piccole e piccolissime imprese di regola insieme col capo lavoravano garzoni privi di qualifiche e d'istruzione, oggi lavorano persone spesso assimilabili a tecnici" (Sylos Labini, 1986).

Si è tentato di introdurre un elemento di misurazione degli effetti determinati dalla riduzione dell'occupazione industriale manifatturiera calcolando un indice sintetico che rapporta tale quota di occupazione alla corrispondente quota di popolazione residente, in Piemonte e nell'Area Metropolitana Torinese (IRES del Piemonte, 1986):

TABELLA 5 - Indice di industrializzazione manifatturiera della popolazione

ANNO	PIEMONTE	AREA METROPOLITANA TORINESE	DIFFERENZA
1982	1.65	1.85	-0.20
1983	1.59	1.75	-0.16
1984	1.46	1.55	-0.09
1985	1.40	1.50	-0.10

Tali indici presentano valori decrescenti, fra il 1982 e il 1985, tanto nell'A.M.T. quanto in Piemonte. La riduzione risulta più ampia nel caso dell'A.M.T.

Il divario fra l'indice regionale e quello torinese si riduce della metà nel periodo, confermando l'effetto di riequilibrio territoriale del peso dell'industria manifatturiera che il riassetto dell'apparato produttivo determina come corollario della concentrazione del ridimensionamento occupazionale operato.

Va da sé che sarebbe utile disporre di informazioni dettagliate sui flussi, utilizzabili per analizzare più a fondo tali questioni e, in particolare, l'intreccio fra le tendenze alla ridislocazione territoriale delle attività e della popolazione. Si può rilevare infatti uno stretto parallelismo fra riduzione degli addetti manifatturieri e riduzione degli abitanti, che pure si manifesta con maggiore intensità nell'area torinese: nel quinquennio 1980-1985 i residenti in Piemonte si riducono di 123.511 unità, di cui ben 107.880 nel comune capoluogo.

Osservando, sia pur sommariamente, i dati demografici relativi alla dinamica 1980-1985 si evidenziano ulteriori elementi di corrispondenza:

TABELLA 6 - Residenti in Piemonte per classi di ampiezza demografica dei comuni 1980-1985

CLASSI COMUNALI	RESIDENTI		$\Delta\%$	DIFFERENZA
	1980	1985		
fino a 5.000 residenti	1.317.040	1.303.646	-1.0	-13.394
5.001-25.000	980.678	990.029	1.0	9.351
oltre 25.000 residenti	2.220.105	2.100.637	-5.4	-119.468
- di cui Torino	(1.143.263)	(1.035.383)	(-9.4)	(-107.880)
TOTALE PIEMONTE	4.517.823	4.394.312	-2.7	-123.511

Alla riduzione dei posti di lavoro corrisponde una riduzione dei residenti e alla riduzione della concentrazione degli addetti corrisponde una riduzione della concentrazione degli abitanti.

Il cambiamento appare dunque rilevante anche rispetto ai processi di dislocazione e ridislocazione della popolazione sia rispetto al territorio regionale (con flusso netto negativo -2,7% nella prima metà degli anni '80), sia rispetto alle tipologie comunali, con un effetto di rivalutazione residenziale limitata alle dimensioni urbane medio-piccole. In altri termini si manifesta un arresto e una inversione del processo di polarizzazione metropolitana (ma non urbana) della popolazione che può produrre fenomeni di diffusione del modello di vita urbano attraverso l'innescò di un processo di "controurbanizzazione" ovvero di deconcentrazione degli abitanti e insieme di rivitalizzazione dei centri urbani medio-piccoli (Petsimeris, 1987).

La connessione fra industrializzazione e concentrazione urbana osservata in passato, nel periodo di più intensa crescita dell'attività produttiva, si conferma anche nella fase di ristrutturazione e riassetto dell'industria attraverso un corrispondente processo di declino o ridimensionamento metropolitano. Tale processo, ancora una volta, produce effetti di maggiore flessibilità residenziale, di crescita del grado di diffusione ed articolazione della popolazione sul territorio (ovvero riduzione non solo della densità, ma anche della eterogeneità degli abitanti), di dilatazione dei complessi sistemi relazionali di tipo urbano ad ambiti territoriali più vasti.

Anche in relazione a tali dinamiche, dunque, occorre dotare l'analisi socio-economica di strumenti più duttili e maggiormente orientati a determinare qualità e direzione del cambiamento.

1.1.3 Transizione alla società post-industriale

Tutte le trasformazioni finora qui enunciate, o esaminate separatamente più in dettaglio, possono essere ricondotte ad un

minimo comun denominatore: la transizione alla società postindustriale. Si tratta dell'individuazione di un processo e dell'evocazione di un nuovo assetto cui tale processo tende, capaci di sintetizzare e spiegare il senso delle profonde trasformazioni osservate riconducendole ad unità all'interno di una tendenza coerente di rinnovamento sociale ed economico complessivo.

I caratteri distintivi della società postindustriale possono essere definiti in relazione alla distribuzione dell'occupazione fra i settori dell'attività produttiva di beni e servizi (cfr. fig. 3); oppure in rapporto alle principali caratterizzazioni che assumono tecnologia, produzione e lavoro:

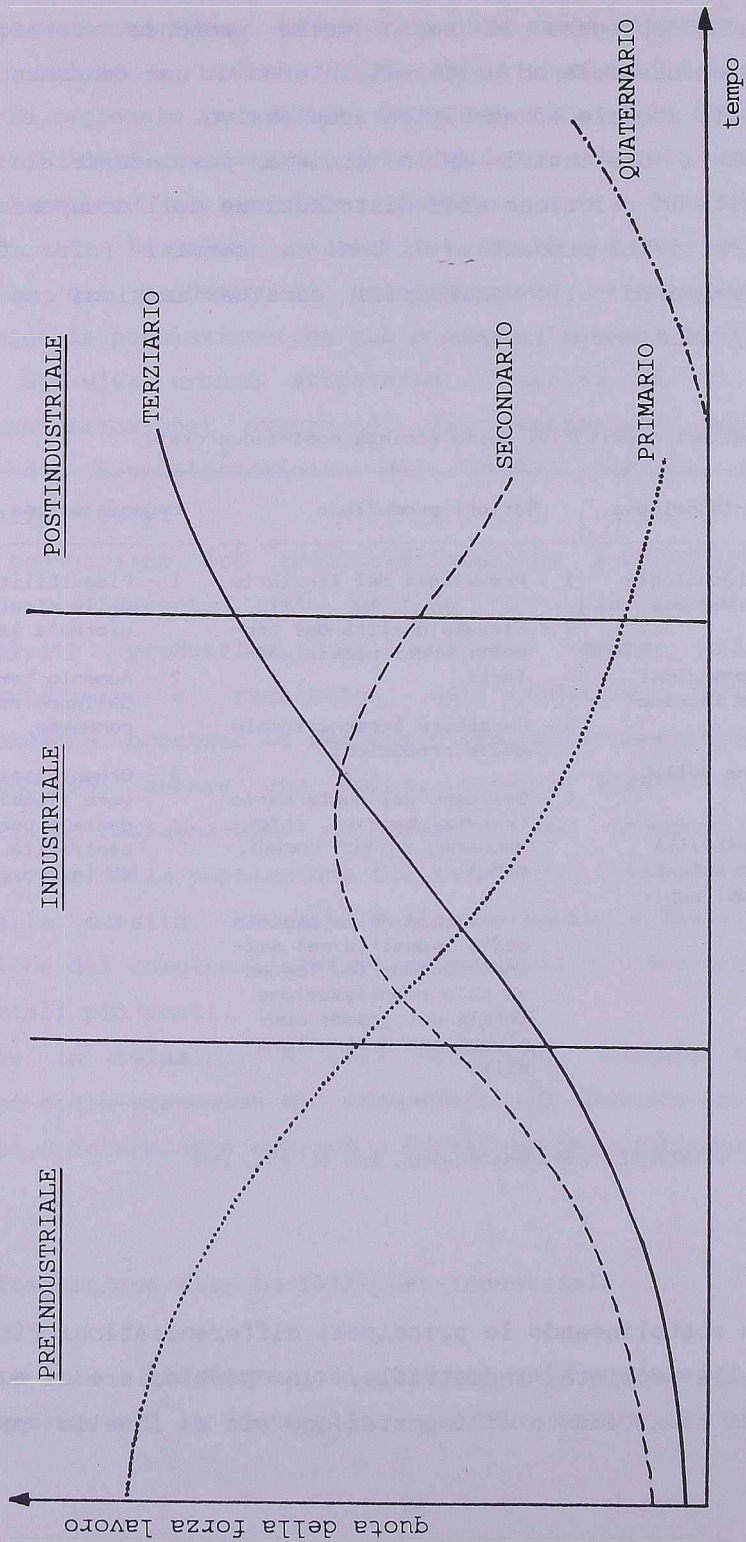
PROSPETTO 1 - Caratteri costitutivi della società post-industriale

livello ricerca-tecnologia	Settori-produzione	Organizzazione del lavoro
1 - Superamento tecnologie rigide della società industriale.	1 - Prevalenza del terziario.	1 - Flessibilità orari e della struttura della giornata lavorativa
2 - Rapidità informazioni su conoscenze tecnico-scientifiche.	2 - Elevata qualità dei prodotti e dei servizi offerti.	2 - Aumento tempo libero. Sviluppo delle attività connesse.
3 - Programmazione della innovazione	3 - Carattere internazionale delle produzioni.	3 - Organizzazione del lavoro flessibile. Tendenza a una perdita di centralità del luogo di lavoro.
4 - Rapida disponibilità delle risorse scientifiche per scopi applicativi	4 - Sviluppo del quaternario (servizi bancari, informazione, servizi sociali, etc.).	
	5 - Prevalenza della piccola media industria nel settore secondario. Tendenza alla robotizzazione totale dei grandi complessi non destrutturabili.	

tratto da: BRUNETTA R. e VENTURINI A., Economia e lavoro, anno XX, n° 1, 1986

oppure ancora sottolineando le principali differenziazioni rispetto ai caratteri della società industriale, in particolare di quelli che evidenziano un ribaltamento di impostazione sia a livello produttivo

Figura 3 - Modello generale settoriale.



tratto da: P.Haggett, Geography. A Modern Synthesis, HARPER e Ron, N.Y., 1979.

sia a livello socio-economico:

PROSPETTO 2

	Caratteri soc. industriale	Caratteri soc. post-industriale
Livello produtt. e di mercato	1 - Standardizzazione produttiva anche con supporti collaterali (mass-media, cultura).	1 - Flessibilità + personalizzazione delle produzioni
	2 - Elevate dimensioni industr.	2 - Piccola dimensione e riduzione peso grande impresa.
	3 - Concentrazione produttiva e urbana.	3 - Diffusione unità produttive e insediamenti sul territorio.
Livello socio-economico	4 - Sincronizzazione e rigidità tempo di lavoro e tempo sociale.	4 - Organizzazione del lav. flessibile + tempi gestiti secondo esigenze individuali.
	5 - Centralizzazione.	5 - Perdita centralità posto di lavoro.

tratto da: BRUNETTA R. e VENTURINI A., idem, 1986.

Un elemento di definizione fondamentale viene individuato nella prevalenza del settore terziario sia in termini di quota di assorbimento occupazionale, sia in termini di quota di fatturato sul prodotto nazionale lordo, rispetto al settore secondario. Ciò è necessario, ma non sufficiente: occorre inoltre che, all'interno del settore terziario, si sviluppino quei servizi di trattamento e diffusione dell'informazione (quaternario) e che, unitamente ai servizi di istruzione, sanità e sociali, assumano un peso ed un ruolo prevalente rispetto ai servizi di tipo tradizionale (intermediazione commerciale).

In Piemonte, rispetto a tutto ciò, si può dire che, con gli anni '80, si entra nella parte di transizione alla società postindustriale. Intanto il terziario diventa il primo settore per quota occupazionale nella prima metà degli anni '80 e tale processo di rafforzamento

continua:

TABELLA 7 - Composizione % dell'occupazione per settori in Piemonte.

ANNI	TERZIARIO	INDUSTRIA	AGRICOLTURA
1980	40.8	48.4	10.8
1985	49.0	41.9	9.1
1986	49.7	41.5	8.8

Dato questo, il fatto che la situazione piemontese si collochi ancora in uno stadio iniziale nella fase di transizione può essere evidenziato in base a due ordini di considerazioni. In primo luogo la composizione interna dell'occupazione terziaria evidenzia l'ancora notevole incidenza dei settori tradizionali: il solo settore commerciale pesa per il 40,5%, mentre nel comparto credito e assicurazioni lavora soltanto l'8% degli addetti terziari:

TABELLA 8 - Composizione dell'occupazione nel terziario al 1985 in Piemonte

Comparti	occupati (000)	composizione %
attività commerciali	349	40.5
trasporti e comunicazioni	82	9.6
credito e assicurazioni	69	8.0
pubblica amministr. e altri	361	41.9
Totale Terziario	861	100.0

In secondo luogo l'importanza del settore terziario in termini di reddito prodotto non sembra ancora corrispondere all'incidenza

occupazionale raggiunta in quanto si nota il persistere di una corrispondenza diretta fra peso percentuale di occupati nell'industria e livello del reddito pro-capite a scala provinciale (con la sola eccezione del caso di Alessandria ove peraltro le attività artigianali - imprese con meno di 10 addetti - rappresentano una quota rilevante delle attività industriali manifatturiere):

TABELLA 9 - Occupazione industriale e reddito pro-capite per provincia in Piemonte

Province	% occupati nell'industria (1986)	reddito pro-capite (Italia=100, 1984)
Vercelli	48.13	131.3
Novara	45.65	125.7
Torino	44.82	120.5
Alessandria	32.57	112.7
Cuneo	32.36	115.0
Asti	30.34	114.4

Gli altri elementi che caratterizzano la società postindustriale e la fase di transizione ad essa, sintetizzati nei prospetti 1 e 2 allegati, si possono tutti riconoscere come agenti o primi effetti del processo di trasformazione economica e sociale cui si è fatto fin qui riferimento. Si tratta dell'innovazione tecnologica, della flessibilizzazione dei sistemi produttivi e dell'organizzazione del lavoro, della discriminante qualitativa che seleziona i nuovi prodotti e i nuovi servizi, della crescita del grado di indifferenza localizzativa e quindi del superamento della concentrazione produttiva e della centralizzazione urbano-metropolitana.

La conseguenza più importante che deriva da tutto ciò è la tendenza ad una "destrutturazione del tempo nel ciclo di vita dell'individuo che non sarà più costretto nella implacabile sequenza di periodi di istruzione, lavoro e pensionamento" (BRUNETTA e VENTURINI, 1986). Ed è proprio in relazione a tale prospettiva che si può ipotizzare una crescente centralità dell'analisi dei flussi nella

ricerca socio-economica, in particolare per la prevedibile crescente alternanza tra periodi di lavoro e periodi di formazione e inoltre per l'agire di una tendenza alla differenziazione all'interno dell'area del lavoro che si manifesta attraverso una articolazione modulare dei contratti rispetto al tempo di lavoro (tempo flessibile, tempo definito, part-time con diverse modulazioni e non solo come metà tempo).

1.2 Nuovi strumenti d'indagine: ridefinizione e riorientamento

Il mercato del lavoro è il principale ambito in cui il processo di trasformazione scarica i suoi effetti: un punto fermo è la necessità di ridefinire e riorientare gli strumenti di conoscenza in rapporto alle nuove condizioni e prospettive.

Occorre modificare ed integrare gli attuali sistemi di rilevazione, individuare le modalità per ottenere statistiche più disaggregate capaci di cogliere la maggiore articolazione degli stock, in ordine, per esempio, al superamento della rigida dicotomia occupazione-disoccupazione, riconoscere che i fenomeni di entrata-uscita dalla condizione lavorativa sono i meno studiati e che dunque vanno individuati e tenuti in maggior conto non solo i flussi fra lavoro e non lavoro, ma anche quelli fra lavoro e scuola o formazione. Assume maggiore rilevanza il concetto di durata o permanenza di un individuo in una certa condizione rispetto al mercato del lavoro e ciò determina una esigenza conoscitiva aggiuntiva rispetto al problema di definire la collocazione dell'individuo rispetto a tale condizione, problema che di per sé si fa più complesso per l'ampliarsi delle possibilità, per l'individuo, di occupare contemporaneamente una pluralità di posizioni.

Le proposte e gli interventi in merito si muovono lungo due direttrici: da un lato la ridefinizione degli stock, dall'altro lato il riordinamento dell'analisi che passa attraverso, in primo luogo, la rivalutazione delle indagini di tipo longitudinale, atte a cogliere la dinamica intrinseca del mercato dal lato dell'offerta di lavoro, e, in

secondo luogo, il recupero in funzione statistica delle fonti di dati amministrativi, individuati come miniera relativamente poco costosa di informazioni utilizzabili previa opportuna organizzazione metodologica.

Per quanto riguarda la ridefinizione degli stock si può partire dalle sollecitazioni, da più parti e più volte espresse, a migliorare la rilevazione trimestrale sulle forze di lavoro dell'ISTAT, che è chiamata a soddisfare crescenti e diversificati bisogni conoscitivi in quanto si tratta indubitabilmente della principale indagine sistematica che si propone di quantificare e descrivere l'occupazione e la disoccupazione dal lato dell'offerta di lavoro. A questo proposito l'ISTAT sta compiendo uno sforzo non indifferente e impegnando notevoli risorse per migliorare la struttura dell'indagine, le rilevazioni e i modi operativi, i metodi di stima delle principali variabili indagate. I programmi di attività relativi impegnano non solo il Servizio forze di lavoro, ma costituiscono il principale contenuto sia del "Progetto campioni" sia del "Progetto qualità dei dati". Il problema viene anche affrontato esaminando la possibilità di realizzare alcuni approfondimenti dell'indagine e ciò costituisce l'oggetto del progetto di ricerca "Forze di lavoro: disegno dell'indagine e analisi strutturali"; tale progetto si sostanzia nella Convenzione di ricerca fra ISTAT e Dipartimento di Scienze Statistiche dell'Università di Padova.

L'indagine sulle forze di lavoro mantiene e rafforza dunque la sua centralità entro l'indirizzo di ricerca dell'ISTAT tendente a riconsiderare e rivalutare tutto il sistema di rilevazioni sulle famiglie, finalizzato anche a realizzare progressivamente un sistema informativo sul lavoro. Lo sforzo di revisione e ammodernamento tende proprio a superare l'incerta definizione degli stock, per esempio nella suddivisione della forza lavoro fra "occupati" e "in cerca di occupazione", tenendo conto di alcuni nuovi concetti introdotti quali la distinzione fra popolazione abitualmente ovvero attualmente attiva (BIT), il concetto di "inoccupazione" o quello di "vuoto occupazionale" (CIRAVEGNA, 1987).

E' interessante, in particolare, la ridefinizione della

disoccupazione in termini di "vuoto occupazionale", inteso come sottoinsieme capace di ricomprendere tutte le forme di sottoutilizzazione di forza lavoro atta ad essere occupata in attività lavorativa retribuita. La formulazione di questo concetto muove dalla considerazione che il termine disoccupazione può assumere una intera gamma di significati in rapporto a diversi contesti specifici di riferimento e ciò sia a livello micro sia a livello macro: i principali individuati e più frequentemente riconoscibili di tali significati possono essere riportati nello schema riassuntivo ed esemplificativo seguente:

DISOCCUPAZIONE

LIVELLO MICRO

LIVELLO MACRO

- | | |
|--|--|
| <ul style="list-style-type: none"> - condizione giuridica - condizione economica - attività - attitudine - stato di necessità | <ul style="list-style-type: none"> - squilibrio domanda-offerta - indicatore squilibrio significativo fra dinamica salari reali e/o monetari - situazione incompleto utilizzo delle potenzialità economiche - indicatore disagio sociale |
|--|--|
-

Di qui, attraverso il concetto di "vuoto occupazionale" si può arrivare ad una ridefinizione dello stock disoccupazione a partire dal riconoscimento di una sua prima e fondamentale ripartizione in sottoccupazione e inoccupazione. Una volta definiti gli occupati in senso pieno come coloro che svolgono un'attività a tempo indeterminato, oppure determinato ma per un periodo superiore ad un anno, si possono individuare i sottoccupati, da intendersi tali sia in senso orario sia in senso produttivo, e ulteriormente ripartibili in retribuiti oppure non retribuiti, in base alla classificazione

riportata nello schema seguente:

SOTTOCCUPATI:

- IN SENSO ORARIO E IN SENSO PRODUTTIVO
 - RETRIBUITI E NON RETRIBUITI
-

- a tempo determinato (periodo inferiore ad un anno)
 - stagionali occupati
 - apprendisti
 - coadiuvanti famigliari
 - lavoratori a domicilio
 - part-time volontario
 - part-time involontario
 - a tempo parziale rispetto alla norma contrattuale
 - occupati a livelli inferiori rispetto a quelli ricopribili (professionali e di qualificazione).
-

Il restante sottoinsieme degli inoccupati, siano essi retribuiti o non retribuiti, rilevati o nascosti, può essere descritto attraverso una schematizzazione basata sulla ulteriore distinzione fra soggetti

INOCCUPATI:

- RETRIBUITI O NON RETRIBUITI
 - RILEVATI O NASCOSTI
-

CON POSTO DI LAVORO

SENZA POSTO DI LAVORO

-
- | | |
|------------------------------------|--|
| - in attesa di iniziare l'attività | - già occupati |
| - sospesi totali | - in cerca di 1 ^a occupazione |
| - stagionali non al lavoro | - offerenti potenziali (lavoratori scoraggiati). |
-

L'articolato schema complessivo che ne risulta, interpretativo e descritto insieme, appare adeguato alle nuove esigenze di analisi ma contemporaneamente rispecchia l'accresciuta complessità che caratterizza la situazione attuale. La riduzione di tale elevato grado di complessità passa attraverso il riorientamento dell'analisi in senso dinamico e in funzione della determinazione dell'entità e della

direzione dei principali flussi in atto. Ciò consente di selezionare via via il campo d'indagine a quei sottoinsiemi e a quegli aspetti che risultano determinanti rispetto alle tendenze prevalenti e inoltre, più in generale, quando le trasformazioni sono di rilevante entità oppure i processi subiscono una accelerazione molto elevata, cresce l'interesse a conoscere le variazioni marginali oltre e più di quelle complessive di periodo in quanto al crescere della durata del periodo si riduce, come si è visto, il numero dei passaggi percepiti e ciò costituirebbe un limite non indifferente alla comprensione della natura dei cambiamenti e dei fenomeni osservati.

Per quanto riguarda più direttamente il secondo elemento caratterizzante i nuovi strumenti d'indagine, ovvero il riorientamento dell'analisi, si può ritenere centrale la rivalutazione delle ricerche di tipo longitudinale.

L'obiettivo è quello di studiare i fenomeni con particolare attenzione all'evoluzione dei loro caratteri nel tempo. Le principali modalità d'indagine utilizzabili a questo scopo sono il riferimento a dati di fonte amministrativa, che forniscono informazioni su soggetti mantenendo via via l'univocità del riferimento individuale; le analisi retrospettive, che si propongono di ricostruire le "storie di vita" individuale e ricomporre poi il quadro complessivo delle trasformazioni così rilevate; l'effettuazione di indagini successive sugli stessi soggetti: si tratta in genere di indagini campionarie su panel di popolazione scelti sia con il criterio della loro rappresentatività generale sia per la particolarità di condizioni che rappresentano (e in questo caso si tratta a volte di piccoli universi: per es. i 7.500 lavoratori Fiat in Cigs posti in mobilità).

A quest'ultima modalità si può ricondurre l'indagine trimestrale dell'ISTAT sulle forze di lavoro, o meglio quella parte di indagine longitudinale che si sta e si intende ampliare e rivalutare: ciò è reso possibile, come si approfondirà in seguito, dalla particolare struttura del campione utilizzato ove, per il tipo di rotazione delle interviste, il 50% nelle famiglie è comune sia in due rilevazioni successive sia in due rilevazioni a distanza di un anno.

La tendenza a riorientare l'analisi in funzione dei nuovi bisogni

conoscitivi produce anche effetti in negativo, di abbandono di alcune impostazioni tradizionali. E' il caso delle informazioni ricavabili dai dati di iscrizione al Collocamento che appaiono sempre meno adatte a descrivere e capire il fenomeno disoccupazione o, più in generale, a fornire indicazioni attendibili circa le dinamiche nel mercato del lavoro. E' sufficiente, a questo proposito, riportare alcune riflessioni espresse a margine di una ricerca della disoccupazione giovanile a Napoli e in Campania avviata nel 1981 (BECCHI COLLIDA', 1986): "Contare i disoccupati a partire da quanti si iscrivono al Collocamento è usuale. E lo è anche quando i conti del Collocamento danno risultati strabilianti rispetto ai dati forniti da altre fonti ufficiali, come accade a Napoli e in Campania, dove -si sostiene- gli iscritti al Collocamento sono 700.000, mentre le indagini ISTAT sulle forze di lavoro registrano, per il 1985, circa 300.000 disoccupati, giovani e non. Uno squilibrio così forte non può non risultare allarmante. Ma vi è di più: al Collocamento non si iscrivono in genere coloro che cercano un lavoro, non avendo nessuna occupazione, ma si iscrivono alcuni di coloro che, occupati o non che siano, ritengono che l'iscrizione sia una chance in più per giungere al posto di lavoro desiderato".

Il senso del riportare questo breve commento sta anche nella sua significatività rispetto all'esigenza di approntare nuovi strumenti d'indagine attraverso la ridefinizione delle variabili e il riorientamento dell'analisi: viene ben esemplificato qui come al mutare dei contenuti e delle caratteristiche dei fenomeni si ha una corrispondente variazione di significato dei dati che occorre ridefinire e riclassificare non solo per se stessi, ma all'interno del processo di riorientamento dell'analisi rispetto alle nuove forme e ai nuovi contenuti delle trasformazioni socio-economiche.

2. LE METODOLOGIE E LE ESPERIENZE

Le metodologie per stimare i flussi, ovvero per ottenere ogni tipo di informazioni organizzate in termini di flusso, si possono ripartire in due fondamentali percorsi di determinazione del dato: quello induttivo e quello deduttivo.

Viene inteso e definito come induttivo quel percorso che, partendo da un campione di osservazioni, individua entità e direzione dei cambiamenti che avvengono in un determinato periodo a livello elementare e, una volta ricostruito il quadro delle trasformazioni in tale ambito, riporta all'universo, per estensione, il contenuto dell'informazione giungendo così a definire i flussi lordi complessivi. E' il metodo, il percorso seguito dall'ISTAT per la costruzione delle matrici di flusso delle forze di lavoro a livello nazionale.

Si può viceversa intendere e definire come deduttivo quel percorso basato su procedimenti di stima elaborativa di dati di stock, in genere molto dettagliati e disaggregati finemente oltre che disponibili a due date successive. Si tratta di un procedimento che, a partire dalle trasformazioni complessive che si sono verificate in un periodo, espresse in termini di flussi netti, ovvero variazioni di stock, attraverso una scomposizione di tali entità in sottoinsiemi elementari (settoriali e territoriali per es.) giunge, con elaborazioni basate su algoritmi di calcolo molto complessi che necessitano di tecnologie informatiche di trattamento dati, a stimare i flussi lordi elementari e, per questa via, a ricostruire il quadro generale della direzione ed entità del cambiamento. E' il metodo, il percorso seguito nei recenti tentativi di analisi tecnico-scientifica dei dati elettorali nell'intento di individuare i flussi lordi che, ad un tempo, sottendono e determinano le variazioni fra due successivi responsi delle urne.

La scelta fra le due modalità spesso è obbligata, essendo determinata a priori dalla natura e dalle caratteristiche delle informazioni di base disponibili per l'elaborazione. Ad esempio, non è

possibile pensare di utilizzare il primo metodo (quello definito come induttivo) nell'analisi dei flussi elettorali perché non si possono organizzare le informazioni elementari in un file longitudinale, ovvero non è possibile ordinare univocamente i dati corrispondenti rilevati al tempo iniziale e al tempo finale.

Analogamente, anche l'analisi della dinamica del mercato del lavoro non avrebbe che il secondo metodo (deduttivo) a disposizione per il trattamento delle informazioni tratte dalla rilevazione trimestrale ISTAT sulle forze di lavoro, se tale rilevazione non fosse consegnata in modo da consentire, almeno in parte, l'accoppiamento dei dati individuali in due successive rilevazioni.

Risulterebbe viceversa possibile, almeno in via teorica, l'applicazione delle metodologie utilizzate nell'analisi del voto e dei flussi elettorali per stimare i flussi della forza lavoro, all'unica condizione di disporre di dati di stock sufficientemente disaggregati: è certamente il caso dei dati censuari, con un limite non indifferente costituito dalla elevata ampiezza del periodo di tempo considerato; potrebbe essere il caso dei dati trimestrali ISTAT sulle forze di lavoro, alla ulteriore condizione di verificare che la disaggregazione richiesta non riduca troppo sensibilmente la significatività del campione.

Le altre sperimentazioni ed esperienze di analisi della dinamica del mercato del lavoro e, più in generale, della evoluzione sociale ed economica, si basano in genere su dati di fonte amministrativa che sono, per loro stessa natura, files di tipo longitudinale e che perciò sono più agevolmente trattabili ed elaborabili con il metodo induttivo. Il principale limite e le difficoltà di quest'ultimo tipo di approccio consistono nella trasformazione del dato amministrativo in dato statistico, operazione non sempre agevole che spesso comporta di dover scegliere un elevato grado di manipolazione dell'informazione di base che complica ed appesantisce il lavoro di indagine nella fase di trattamento dei dati, ma migliora il grado di attendibilità e rappresentatività dei risultati. Oppure si può optare per una impostazione dell'analisi più aderente alla forma e struttura del dato amministrativo di riferimento, che consente di riutilizzare

successivamente le metodologie di elaborazione dati adottate, ma spesso ciò comporta di limitare fortemente i risultati ottenuti dal punto di vista della loro capacità ed incisività interpretativa oppure il grado di flessibilità del loro utilizzo.

I due esempi di esperienze di indagini in termini di flusso svolte in Piemonte cui si è già accennato, la ricerca sulla demografia industriale (CONTINI, ASCOLI, REVELLI, 1986) e quella sui flussi occupazionali e localizzativi dell'industria manifatturiera (IRES del Piemonte, 1987), sono esemplari rispetto a questa dualità di approccio al dato di fonte amministrativa, come si vedrà in seguito, nel presente capitolo, esaminandone i contenuti metodologici.

2.1 I flussi di forza lavoro

La procedura utilizzata dall'ISTAT per abbinare i dati individuali di rilevazioni successive e pervenire quindi alla determinazione delle matrici di transizione trimestrali o annuali, che contengono appunto i dati che costituiscono l'obiettivo di tale procedimento di calcolo, è stata descritta in modo sintetico ed esauriente (MORIANI, 1981) nel Supplemento al Bollettino mensile di Statistica n.15 del 1981 e qui riportata in Appendice.

I dati relativi ai movimenti della popolazione dall'una all'altra condizione definita in base alla classificazione della forza di lavoro, costituiscono l'oggetto di una pubblicazione dell'ISTAT che si può ormai considerare regolare, articolata e consolidata rispetto alle prime sperimentazioni impostate fin dal 1971 e riprese, su nuove basi metodologiche, nel 1981.

Le informazioni di base utilizzate sono quelle ricavabili dalle rilevazioni trimestrali delle forze di lavoro. Tali osservazioni elementari vengono selezionate, organizzate ed elaborate in modo da ottenere la quantificazione dei flussi lordi, in entrata e in uscita, ovvero definire l'entità dei passaggi, in un dato periodo di tempo, dall'uno all'altro settore di attività (agricoltura, industria, altre attività terziarie, pubblica amministrazione) per quanto riguarda la

classificazione dell'universo degli occupati e, più in generale, dall'una all'altra condizione di appartenenza all'insieme delle forze di lavoro (occupati, in cerca di occupazione) o della popolazione presente (forza di lavoro, non forze di lavoro in età lavorativa o non).

Gli intervalli temporali di riferimento considerati sono il trimestre e l'anno: i flussi trimestrali registrano però i movimenti che si manifestano tra due successive rilevazioni effettuate in momenti precisi; i flussi annuali sono invece determinati come risultante del confronto fra le consistenze medie, rispetto alle quattro rilevazioni annuali, riferite a due anni successivi.

Le consistenze utili per l'elaborazione dei dati in funzione della determinazione dei flussi sono quelle del sub-campione (pari al 50% lordo) rispetto all'universo delle forze di lavoro, costituito dalle informazioni elementari accoppiabili, ovvero riferite allo stesso soggetto e relative a due successive rilevazioni (flussi trimestrali) o a due gruppi corrispondenti di rilevazioni (flussi annuali).

Il livello territoriale d'indagine è quello nazionale.

Gli scopi di una tale impostazione di ricerca e di elaborazione dei dati sono molteplici: arricchimento degli strumenti conoscitivi; possibilità di utilizzare i risultati come inputs per modelli interpretativi o previsivi delle dinamiche del mercato del lavoro; verifica continua sulla efficienza stessa delle rilevazioni su cui si basano.

Il metodo adottato è riconducibile al criterio delle indagini longitudinali che consentono il confronto delle risposte fornite dagli stessi individui alle stesse domande poste in tempi diversi. Ciò è possibile per l'adozione di un criterio di rotazione delle famiglie nel campione di rilevazione delle forze lavoro: tale sistema di rotazione è stato introdotto nel 1971.

La rotazione funziona in modo tale che ogni gruppo o sezione delle famiglie che compongono il campione viene intervistato in un trimestre, in quello successivo e nel primo trimestre corrispondente dell'anno seguente: poiché le sezioni interessate all'indagine in ciascun trimestre sono quattro, ne risulta un grado di accoppiabilità

dei dati pari al 50% (in line a teorica) del campione complessivo, sia a cadenza trimestrale sia a cadenza annuale (cfr. schema in Appendice). In pratica la misura indicata si riduce per due ordini di motivi. Il primo comprende tutte le operazioni di sostituzione delle famiglie intervistate che debbono essere effettuate per varie cause: trasferimenti, rifiuti, irreperibilità nella settimana di riferimento, ecc.. Il secondo è legato alla doppia stratificazione del campione che comporta la sostituzione, con cadenza annuale, anche di una parte dei comuni con meno di 20mila abitanti: ciò produce una riduzione del grado di accoppiabilità dei dati individuali fino al 40% circa del campione complessivo.

L'accoppiamento dei dati individuali viene ottenuto attraverso una procedura a due fasi: accoppiamento delle famiglie, mediante una codificazione che comprende le informazioni su regione, provincia, comune, area di rilevazione, numero progressivo delle famiglie, sezione del campione; accoppiamento degli individui cui si compone la famiglia attraverso una codificazione basata essenzialmente su età e sesso, oltre che, per controllo, relazione col capofamiglia e grado di istruzione, prevedendo naturalmente margini di tolleranza per le variazioni che possono essere state determinate dal tempo trascorso tra la prima e la seconda rilevazione.

Il riporto all'universo del sub-campione degli accoppiabili è l'operazione immediatamente successiva. Consiste nell'individuare innanzitutto l'universo di riferimento del sub-campione in esame, ovvero determinare l'entità della popolazione accoppiabile in base alla formula dell'equazione della dinamica della popolazione:

$$Pa = P1 - N - I = Po - M - E$$

ove Pa indica la popolazione residente accoppiabile; P1 la popolazione residente complessiva al tempo 1; N i nati nel periodo (fra le due rilevazioni considerate); I gli immigrati nel periodo; Po la popolazione residente complessiva al tempo 0; M i morti nel periodo; E gli emigrati nel periodo.

La matrice dei dati grezzi è costruita, come passo successivo

della procedura, determinando i vettori della parte centrale della matrice di transizione attraverso i coefficienti di riporto al sub-universo della popolazione accoppiabile. Tali vettori rappresentano la classificazione degli individui secondo due caratteri socio-economici: quelli posseduti al tempo iniziale 0 e quelli posseduti al tempo finale 1. I vettori di riga individuano la ripartizione degli individui per carattere al tempo 0, ovvero la struttura per provenienza; i vettori di colonna individuano la struttura al tempo 1, ovvero la struttura per destinazione. La matrice ottenuta contiene, a questo punto, la classificazione bivariata secondo origine e destinazione degli individui.

L'ultima fase della procedura consiste nelle operazioni di quadratura della matrice che comportano alcune elaborazioni suppletive per arrivare alla convergenza tra la matrice dei dati grezzi e i dati di cornice ed ottenere così il risultato atteso: la definitiva matrice di transizione. I dati di cornice consistono di due colonne e due righe aggiuntive necessarie per chiudere la matrice: si tratta rispettivamente degli emigrati e dei morti (colonne) e degli immigrati e dei nati (righe) nel periodo considerato. Si tratta di dati esogeni rispetto alla rilevazione della forza di lavoro che vengono determinati in base alle statistiche demografiche correnti.

I primi risultati presentano in genere problemi di divergenza sia all'interno della matrice di dati grezzi, sia in rapporto ai dati di cornice. Tali effetti sono essenzialmente riconducibili a due circostanze: una deriva dal fatto che la prima matrice dei flussi è il risultato di una estensione all'universo a partire dal sub-campione degli individui accoppiabili e dunque diverso dal campione complessivo utilizzato per la rilevazione sulle forze di lavoro; la seconda deriva dalla natura esogena dei dati di cornice che perciò non sempre risultano immediatamente compatibili con le stime della matrice dei dati grezzi.

Per realizzare la convergenza della matrice di transizione definitiva occorre allora annullare le differenze tra i totali della matrice interna e i totali imposti ottenuti mettendo a confronto le due estensioni del campione fra loro e con i dati esogeni relativi

alle entrate e alle uscite dell'universo di riferimento: tale operazione viene eseguita facendo ricorso al metodo RAS (cfr. Appendice 1) che consiste nel riproporzionamento degli elementi di una matrice quadrata in modo da eguagliare le somme di riga e le somme di colonna.

A completamento del quadro informativo vengono infine calcolate, sulla base delle matrici definitive dei valori assoluti, le corrispondenti matrici dei coefficienti orizzontali e verticali.

I coefficienti orizzontali, espressi in termini di distribuzione percentuale di riga, rappresentano i tassi di uscita dalla condizione iniziale, ovvero i flussi secondo l'origine.

I coefficienti verticali, espressi in termini di distribuzione percentuale di colonna, rappresentano i tassi di entrata nella condizione finale, ovvero i flussi secondo la destinazione.

I coefficienti della diagonale principale rappresentano, in entrambe le matrici dei coefficienti, i tassi di permanenza nella stessa condizione nel periodo considerato, ovvero le quote di individui che, per ciascuna condizione, non sono interessati dai movimenti di flusso.

I tassi di uscita, i tassi di entrata e i tassi di permanenza possono inoltre essere interpretati in termini di misura della probabilità di passaggio dall'una all'altra condizione nel periodo, oppure di stabilità nella condizione iniziale; si tratta di un aspetto della teoria delle matrici di transizione su cui si fondano i tentativi di utilizzo dei dati di flusso per giungere ad estrapolazioni previsive di tendenza.

2.1.1 Risultati a scala nazionale

Per illustrare la portata e la capacità interpretativa dei risultati così ottenuti, conviene rifarsi ad un esempio di utilizzo dei dati contenuti nelle matrici di transizione a livello nazionale (MAROTTA, PUGLIESE, 1986) che risulta particolarmente utile a mettere in evidenza il diverso contenuto informativo dei confronti effettuati

in termini di flussi rispetto a quelli riferiti alle variaizoni di stock.

I dati presi in esame si riferiscono ai flussi intersettoriali della forza lavoro (valori medi 1983-1984, sia in valori assoluti sia in valori percentuali) e sono organizzati in una tabella-matrice ove sulle righe si possono leggere le entità dei flussi in uscita rispetto a ciascun settore considerato, sulle colonne i corrispondenti flussi in entrata e sulla diagonale principale si individuano il numero e le quote di individui appartenenti a ciascun aggregato che non sono stati interessati, nel periodo considerato, dai movimenti di flusso.

Al fine di meglio esaltare la funzione esemplificativa, viene qui riportata una sola riga e una sola colonna di dati, quelle riferite al settore industria: la tabella 10 contiene i valori assoluti, le entità dei flussi; la tabella 11 i valori percentuali dei flussi in uscita dal settore industria.

In tabella 10 il totale di riga rappresenta lo stock di occupati nel settore industriale al tempo iniziale (1983); il totale di colonna il corrispondente stock a fine periodo (1984). I dati, espressi in migliaia di unità, sono pari rispettivamente a 5.290,4 e 5.087,4 posti di lavoro. Il confronto di stock ci dice che nel periodo 1983-1984 in Italia si sono perduti 203mila posti di lavoro nell'industria, pari al 3,8% della consistenza iniziale. Il dato della casella posta sulla diagonale principale nella tabella 10 (3.978,5) rappresenta la quota di addetti che, nel periodo considerato, ha mantenuto il posto di lavoro nel settore industria. La differenza fra questo valore e la consistenza iniziale di addetti (5.290,4) definisce l'entità dei flussi in uscita dal settore industria: tale entità è pari a 1.311,9, ovvero oltre un milione e trecentomila addetti occupati nell'industria al 1983 hanno mutato tale loro condizione nel corso dell'anno successivo.

In altri termini ciò significa che l'entità dei flussi è di oltre sei volte superiore rispetto alla entità della variazione di stock. Ciò non toglie, ovviamente, che da un punto di vista sociale sia più rilevante l'entità dei posti di lavoro perduti, ancorché relativamente molto minore del valore complessivo dei cambiamenti di condizione

TABELLA 10 - Flussi interindustriali della forza lavoro, valori medi 1983-1984 a scala nazionale
valori assoluti (000)

SETTORI	AGRICOLTURA	INDUSTRIA	COSTRUZIONI	TERZIARIO PRIVATO	PUBBLICA AMMINISTR.	DISOCCUPA- ZIONE	NON FORZE DI LAVORO	MORTI E EMIGRATI	TOTALE TEMPO 0
Agricoltura		50.2							
Industria	51.8	3978.5	205.9	509.3	125.9	113.8	275.8	29.4	5290.4
Costruzioni		179.3							
Terziario privato		401.4							
Pubblica Amministrazione		105.4							
Disoccupazione		129.8							
Non forze di lavoro		235.8							
Nati e immigrati		7.1							
TOTALE TEMPO 1		5087.4							

TABELLA 11 - Distribuzione percentuale dei flussi in uscita. Valori medi 1983-1984 a scala nazionale

SETTORI	AGRICOLTURA	INDUSTRIA	CONSTRUZIONI	TERZIARIO PRIVATO	PUBBLICA AMMINISTR.	DISOCCUPA- ZIONE	NON FORZE DI LAVORO	MORTI E EMIGRATI	TOTALE TEMPO 0
Agricoltura		1.9							
Industria	0.9	75.2	3.9	9.6	2.3	2.1	5.2	0.5	100.0
Costruzioni		8.7							
Terziario privato		6.2							
Pubblica Amministrazione		2.5							
Disoccupazione		5.7							
Non forze di lavoro		0.7							

entro cui si viene a determinare e di cui rappresenta il saldo netto.

Appare comunque incontestabile l'utilità di conoscere l'ambito complessivo delle trasformazioni e dei cambiamenti nella struttura occupazionale sia perché ciò costituisce un elemento importante per una corretta valutazione delle singole dinamiche di settore, sia per scegliere ed approntare correttamente gli strumenti per affrontare i problemi (la flessione occupazionale) che tali trasformazioni e cambiamenti pongono.

In tabella 11 il valore che si colloca sulla diagonale principale (75,2%) indica il tasso di permanenza degli addetti all'industria nel settore di appartenenza: un addetto su quattro cambia la propria posizione fra il 1983 e il 1984. Gli altri valori di riga indicano la destinazione di tale cambiamento: prevalgono l'occupazione nel terziario privato (il 9,6% del totale addetti 1983) e il passaggio fra le non forze di lavoro (5,2%).

I valori di colonna di tabella 11 indicano la misura dei flussi in uscita dalle altre condizioni che hanno per destinazione l'occupazione industriale che assorbe l'8,7% degli addetti alle costruzioni a inizio periodo, il 6,2% del terziario privato e soltanto il 5,7% del valore complessivo di disoccupazione iniziale.

2.1.2 La scala regionale

L'unica esperienza finora compiuta di costruzione delle matrici dei flussi delle forze di lavoro a scala regionale è quella della Lombardia.

La procedura e i programmi di elaborazione dei dati che sono stati utilizzati sono gli stessi che l'ISTAT impiega a scala nazionale. Le prime sperimentazioni sono state avviate, a cura del Servizio Statistica della Regione Lombardia, non appena è stato realizzato, in tale regione, l'ampliamento del campione di rilevazione della forza di lavoro (1980).

Tale ampliamento è una condizione di base per poter procedere alla elaborazione delle matrici di flusso a scala regionale in quanto il

campione ridotto dei dati utili allo scopo (informazioni accoppiabili) renderebbe, a tale scala, privi di significatività i risultati ottenuti. L'insieme delle informazioni accoppiabili individuate rappresenta infatti, come si è visto, circa il 40% del campione complessivo delle forze di lavoro.

In Piemonte tale ampliamento è stato realizzato ancor prima che in Lombardia (1979), ma ciò al fine di disporre di risultati validi a livello provinciale e per ottenere un grado di significatività del dato regionale analogo a quello del dato nazionale. Il fatto che l'ampliamento del 1980 sia stato subito finalizzato, in Lombardia, anche alla determinazione dei flussi dipende, probabilmente, dalla diversa e maggiore sensibilità che lì si avverte per i fenomeni di mobilità sociale, occupazionale ed economica, in relazione alla maggiore articolazione che ha caratterizzato la società lombarda rispetto a quella piemontese, quantomeno fino al 1980 in termini apprezzabili.

Alcuni aggiustamenti dei programmi di elaborazione si sono resi necessari, nell'esperienza lombarda, per renderli compatibili con la diversa scala territoriale di riferimento. Ad esempio, nell'operazione di riporto all'universo del sub-campione di informazioni accoppiabili, il parametro relativo alla distribuzione regionale del dato viene sostituito con quello della distribuzione provinciale, che peraltro, in Lombardia, risulta sufficientemente articolata. Recentemente si è interrotta, per un periodo, l'attività di elaborazione delle matrici dei flussi essendosi reso necessario, preventivamente, procedere alle operazioni di revisione della serie di dati della rilevazione delle forze di lavoro (retrospettivamente fino al 1980) a causa del nuovo dimensionamento del campione che tiene conto delle risultanze del censimento 1981 nella definizione dell'universo di riferimento.

La possibilità di utilizzo dei risultati, a scala regionale, si arricchiscono ulteriormente, rispetto a quelle già sottolineate per quanto riguarda le matrici nazionali, in quanto disponendo dei dati di flusso a scala regionale è possibile operare anche un confronto fra flussi regionali e flussi nazionali e quindi verificare il grado di convergenza o di divergenza delle trasformazioni che tali informazioni

descrivano.

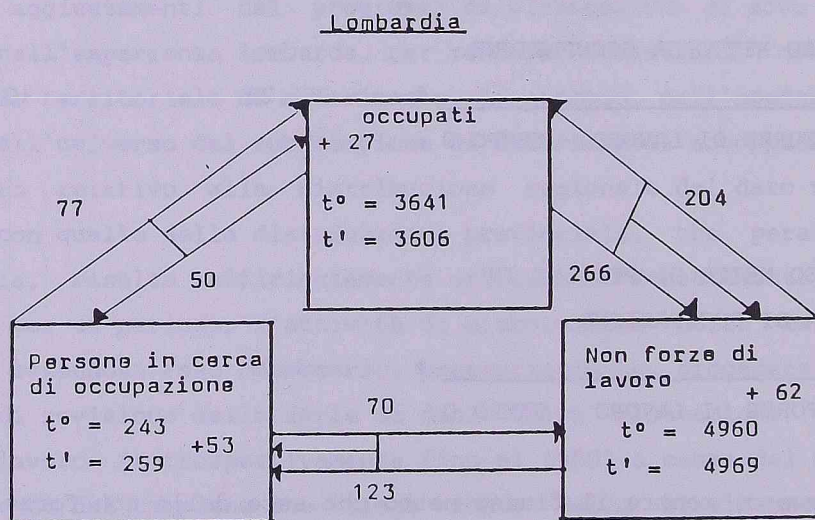
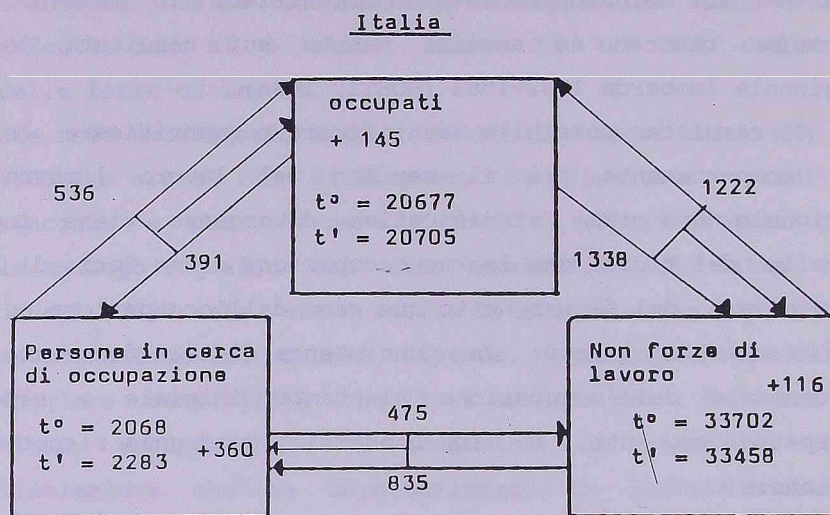
Anche qui è possibile esemplificare la potenzialità interpretativa dei dati sul mercato del lavoro organizzati in termini di flusso, facendo ricorso ad analisi basate sui risultati della matrice regionale lombarda 1982-1983 (ROSTI, 1986).

E' risultato possibile identificare e quantificare le differenze di comportamento tra il mercato del lavoro lombardo e quello nazionale. Una prima significativa divergenza viene individuata a livello dei flussi che legano occupazione e non forze di lavoro e, in particolare, del flusso netto che esce dall'occupazione ed alimenta le non forze di lavoro: la consistenza di tale fenomeno, espressa in percentuale dell'occupazione al tempo iniziale e riportata nel prospetto seguente, è, in Lombardia, quadrupla rispetto alla media nazionale:

INDICATORI	LOMBARDIA	ITALIA
FLUSSO NETTO DA OCCUPAZIONE		
<hr/>	% +1,26	+0,34
NON FORZE DI LAVORO - TEMPO 0		
FLUSSO NETTO DA PERSONE IN CERCA DI OCCUPAZIONE		
<hr/>	% -1,07	-1,07
NON FORZE DI LAVORO - TEMPO 0		

Ovvero, mentre il flusso netto che esce dalle non forze di lavoro ed alimenta le persone in cerca di occupazione è perfettamente identico, in senso relativo, in Lombardia e in Italia (-1,07%), il flusso netto da occupazione a non forze di lavoro è, in Lombardia, quattro volte più elevato rispetto al dato nazionale (1,26% contro 0,34%). Tale considerazione è possibile ricavare una volta ricostruito il prospetto dei flussi netti fra occupazione, non forze di lavoro e persone in cerca di occupazione così come viene riportato in figura 4.

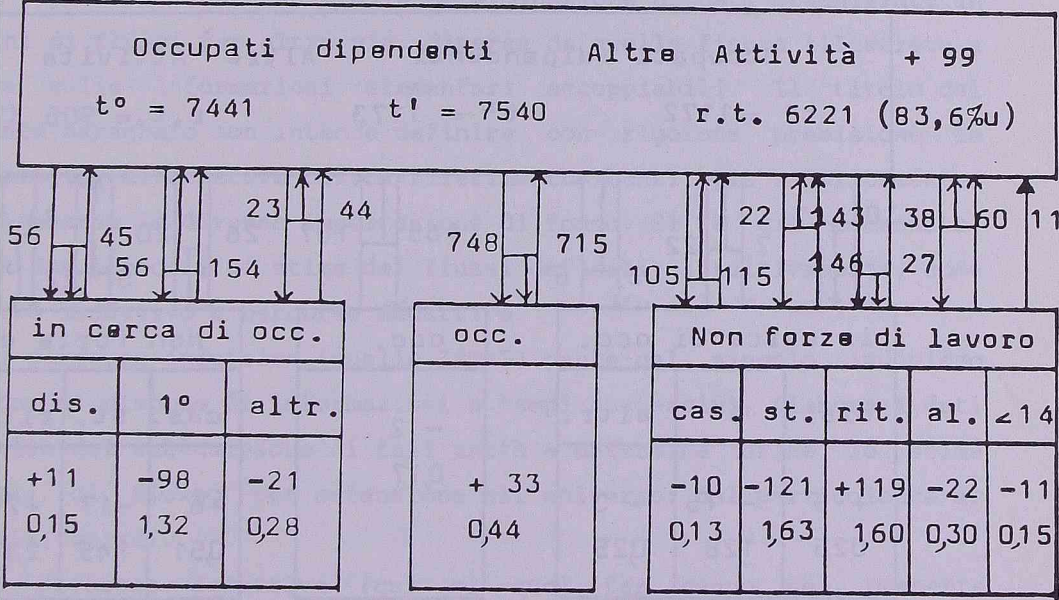
Figura 4 - Flussi in entrata, in uscita e netti tra occupati, persone in cerca di occupazione e non forze di lavoro in Lombardia e in Italia, 1982-1983, media annua, v.a. migliaia.



tratto da: L.Rosti, L'analisi dei flussi sul mercato del lavoro in Lombardia e in Italia, paper presentato all VII Conferenza Italiana di Scienze Regionali, Urbino, settembre 1986.

Figura 5 - Flussi lordi e netti tra occupazione terziaria dipendente e altra occupazione, persone in cerca di occupazione e non forze di lavoro. Media 1982-1983, v.a. migliaia e % su occupazione terziaria dipendente al tempo t°.

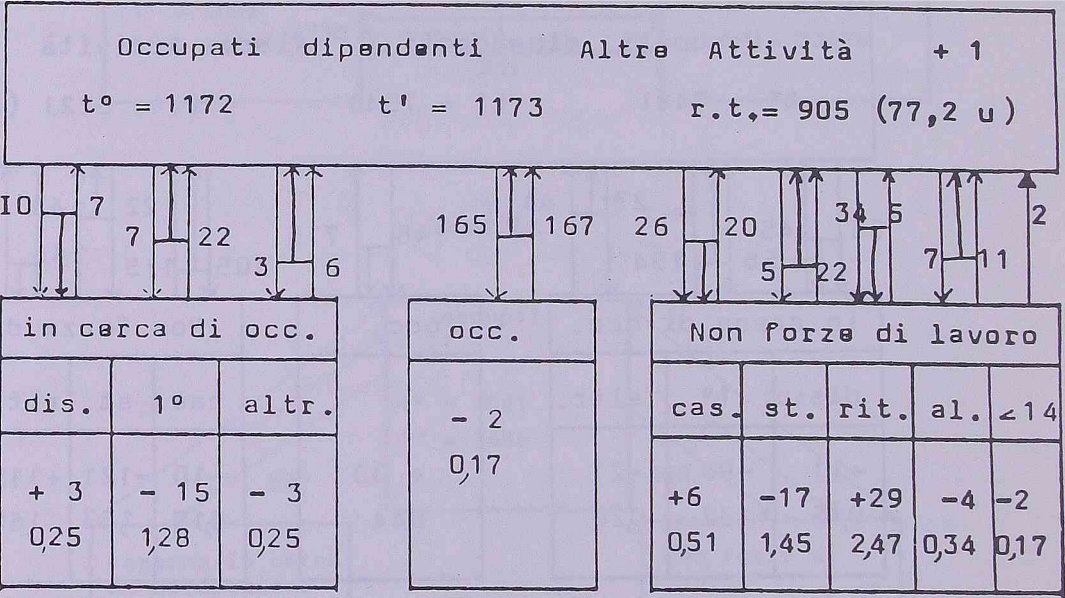
Italia



tratto da: cfr. figura 4.

Figura 6 - Flussi lordi e netti tra occupazione terziaria dipendente e altra occupazione, persone in cerca di occupazione e non forze di lavoro. Media 1982-1983, v.a. migliaia e % su occupazione terziaria dipendente al tempo t°.

Lombardia



tratto da: cfr. figura 4.

Un altro esempio di differenza rilevante Lombardia-Italia che si è potuta individuare si manifesta nel rapporto tra l'occupazione del settore terziario e l'occupazione degli altri settori di attività: mentre in Lombardia l'occupazione terziaria riceve un flusso netto di 2.000 unità dal resto dell'occupazione, in Italia sono gli altri settori di attività ad assorbire ben 33.000 unità lavorative dal settore terziario (cfr. figura 5 e 6).

2.2 I flussi da elaborazione di stock

E' possibile arrivare alla determinazione di dati organizzati in termini di flussi per altra via, diversa da quella finora illustrata e basata sulle informazioni elementari accoppiabili. Il titolo del presente paragrafo non intende definire con rigorosa precisione le metodologie alternative cui fa riferimento, bensì, più semplicemente, sottolinearne la diversa impostazione di fondo. Si è già accennato, qui, a due percorsi di stima dei flussi definiti, intuitivamente, come percorso induttivo e percorso deduttivo.

Il percorso induttivo (quello ISTAT) parte dal singolo individuo per cui si dispone di informazioni a tempi successivi, elabora i dati a livello del sub-campione di tali unità e determina infine le stime generali di flusso per estensione all'universo: dal particolare al generale, quindi.

Il percorso deduttivo (cui si vuol far cenno nel presente paragrafo) parte dai dati aggregati per risalire ad informazioni sul comportamento degli individui operando trasformazioni sulle informazioni che consentano di passare da dati in serie storica a dati sulle transizioni che avvengono tra uno stato e l'altro del sistema: dal generale al particolare, quindi.

Questa seconda metodologia è stata studiata con particolare riferimento ad un tipico caso di applicazione che esclude a priori la possibilità di ricostruire serie storiche di informazioni accoppiabili: si tratta dell'analisi dei dati elettorali (DRAGHI, 1987).

Una recente ed esauriente formulazione scientificamente rigorosa di tale metodologia è riportata nell'Appendice 2 (SCHADEE e CORBETTA, 1984). Può essere di qualche interesse approfondire alcuni aspetti, ovvero la logica generale, la possibilità di generalizzazione della sua applicazione in altri campi di analisi, le ragioni della relativamente recente riscoperta di tale impostazione metodologica per costruire strumenti di analisi per la ricerca sociale.

Il modello di analisi dei dati elettorali utilizza la tecnica della regressione lineare e i metodi sviluppati possono essere applicati a qualsiasi serie di dati aggregati con il solo vincolo della coerenza rispetto alle condizioni formali del modello stesso. Solo recentemente si è assistito ad un forte ritorno di interesse per l'analisi, sociologica in particolare, basata su dati aggregati e ciò in quanto sembrano superati, sul piano metodologico, gli ostacoli che hanno determinato per anni la preferenza per le tecniche di rilevazione dati fondate su interviste e questionari: l'ostacolo principale consisteva nella presunta dimostrazione che i coefficienti di correlazione fra dati aggregati di per sé nulla dicono sui valori dei coefficienti di correlazione esistenti a livello individuale (ROBINSON, 1950).

Una volta appurato che tale critica è infondata, quantomeno nella interpretazione estensiva che se ne fece, ovvero che è impossibile risalire dai dati aggregati ad informazioni sul comportamento degli individui, ridiventa praticabile l'idea che i flussi e la loro entità possano essere individuati come i coefficienti di un modello di regressione lineare ove gli stock rappresentano le variabili indipendenti e le variazioni di stock la variabile dipendente.

In altri termini si tratta di proporsi di scomporre il flusso lordo conosciuto (la variazione di stock nel periodo) nei vari flussi netti che lo determinano.

Anche nello specifico campo di ricerca sul mercato del lavoro si può trovare qualche esempio di posizione del problema in questi termini, ovvero riconoscere che un modo per operare la trasformazione sulle informazioni passando da dati in serie storica a dati sulle transizioni che di volta in volta avvengono tra uno stato e l'altro

del sistema e quindi ai comportamenti individuali è quello di stimare, sulla base dei valori storici forniti dalle statistiche ufficiali, la matrice dei coefficienti di transizione di una "macchina a stati determinati" o "macchina senza input" di Ashby (MICELI e GRAMAGLIA, 1981).

Questo schema di ragionamento può essere ripercorso attraverso una esemplificazione estremamente semplificata: un modello dei flussi di forza lavoro tra la condizione di occupato e quella di non-occupato viene rappresentato secondo la formulazione di FORRESTER (cfr. figura 7); si assume che la ipotesi più ovvia circa la relazione tra variabili di livello e di tasso, sia costituita dalla proporzionalità diretta di ogni tasso con il livello di partenza (freccie tratteggiate) e si fa l'ipotesi che il sistema sia chiuso.

Come si vede, la relazione che si viene ad instaurare tra livelli e tassi è di tipo molto particolare: 1) i flussi in uscita da una condizione sono sempre proporzionali alla dimensione della condizione stessa; 2) i tassi dipendono solo dai livelli (non ci sono variabili ausiliarie), in altri termini il sistema è completamente autoregressivo: lo stato al tempo $t=1$ dipende dallo stato al tempo $t=0$ tramite una certa trasformazione (incremento); 3) una quota fissa di individui passa, in un certo periodo, da un livello all'altro.

Nei termini di Ashby questo sistema corrisponde ad una "macchina senza input" la cui trasformazione anziché essere determinata è stocastica (cfr. figura 8, dove A,B,C,D sono i coefficienti di transizione incogniti). Ove risulti possibile stimare i coefficienti di transizione sulla base dei dati delle serie storiche, risulta per conseguenza possibile simulare dinamicamente il comportamento del nostro sistema, determinato dalla matrice dei coefficienti di transizione.

Le serie storiche di dati potrebbero essere costituite dalle stesse informazioni di stock determinate in base alla rilevazione trimestrale delle forze di lavoro per estensione all'universo del campione complessivo utilizzato.

L'applicazione pratica di queste metodologie di stima dei flussi comporta però una riduzione del grado di distribuzione del dato

Figura 7 - Modello dei flussi occupato-non occupato.

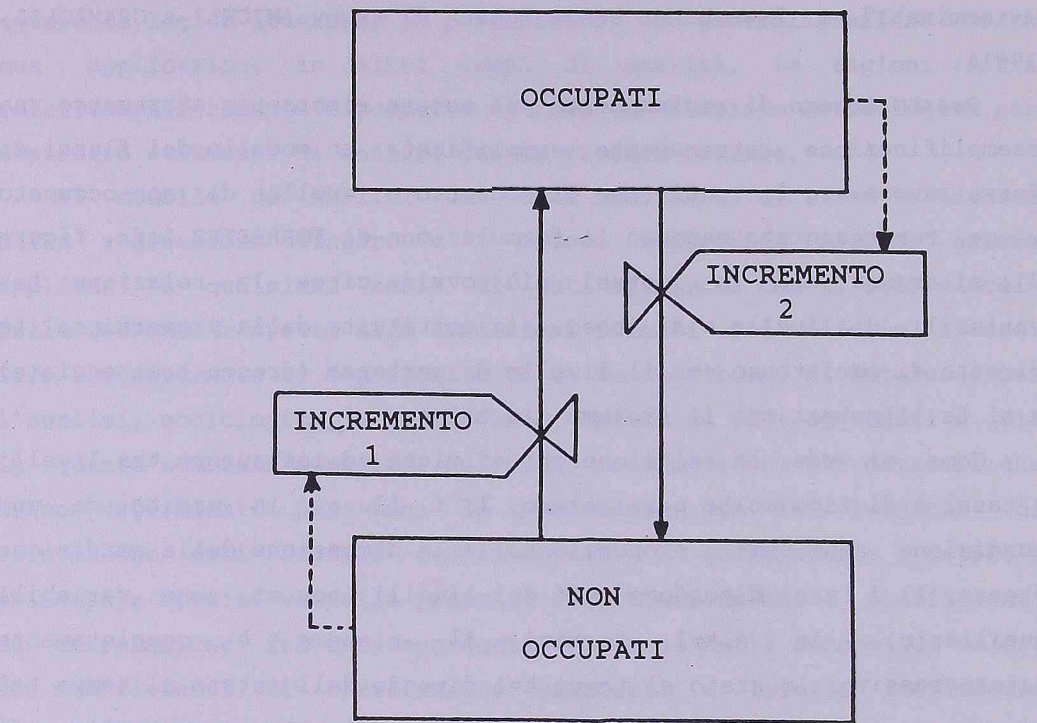


Figura 8 - Matrice dei coefficienti di transizione.

↓	OCCUPATI	NON OCCUPATI
OCCUPATI	A	B
NON OCCUPATI	C	D

(territoriale o settoriale per es.): si può cioè ritenere che ove si disponga di dati di stock a livello comunale (o di classe di attività) si potranno ottenere stime dei flussi soltanto a livello di provincia (o di ramo di attività) perdendo, nel corso della fase di elaborazione, almeno una "qualità" distributiva del dato di partenza.

I maggiori problemi, nell'applicazione di tale metodologia, consistono nella grande complessità degli algoritmi di calcolo che, se è risolvibile attraverso l'utilizzo dei più recenti elaboratori, pone comunque la questione della accessibilità delle informazioni di base in forma direttamente trattabile, ovvero meorizzata su supporti magnetici. Inoltre, se le informazioni di carattere elettorale sono disponibili in forma dettagliata e territorialmente molto articolata, non così è per le informazioni socio-economiche e ciò può costituire un limite rispetto alle possibilità pratiche di generalizzazione applicativa del modello di analisi dei dati elettorali che, a livello teorico-metodologico, non presenta difficoltà.

2.3 Le sperimentazioni su dati amministrativi in Piemonte

Per completare la pur breve e sommaria ricognizione fin qui condotta sulle metodologie adottate nelle esperienze di analisi dell'informazione finalizzata alla scomposizione dei singoli movimenti di cambiamento e delle singole direzioni di tendenza, ci si può ancora brevemente soffermare sui contenuti metodologici di due esperienze già accennate che, con particolare riferimento alla realtà regionale piemontese e alle dinamiche imprenditoriali ed occupazionali del settore industriale, si propongono di giungere ad alcuni risultati, più o meno sistematici, utilizzando come dati di base gli archivi disponibili di fonte amministrativa.

Si tratta dello studio sulla demografia industriale del Piemonte (CONTINI, ASCOLI e REVELLI, 1986) che fa riferimento alle serie di dati ufficiali INPS ed analizza il periodo 1977-1982 e della ricerca sui flussi occupazionali e localizzativi dell'industria manifatturiera piemontese in Provincia di Torino (IRES del Piemonte, 1987) che è

basata sui dati del Registro Ditte delle C.C.I.A.A., integrati dalle informazioni di fonte UPLMO ed altre minori ed analizza il periodo 1980-1985.

Data la natura delle informazioni di base, entrambe queste analisi fanno sostanzialmente riferimento al metodo di accoppiamento delle informazioni via via disponibili riferite alla stessa unità elementare di osservazione. Entrambe, inoltre, scontano la necessità di un forte impegno preventivo di carattere metodologico per operare le necessarie verifiche del dato e renderlo così utilizzabile a fini statistici. Tale impegno risulta particolarmente oneroso anche in relazione al fatto che le metodologie adottate vengono costruite in qualche modo a misura del tipo di informazione che si utilizza e le verifiche di coerenza e di corretta accoppiabilità dei dati debbono essere operate non su un ristretto campione, bensì sull'intero universo, in entrambi i casi in esame costituito da un file di imprese e/o unità locali industriali.

Tutto ciò finisce di apparire, a posteriori, non adeguato ai concreti risultati ottenuti che forniscono indicazioni originali e molto dettagliate, ma circoscritte in un ambito temporale definito e ristretto. In altri termini tali iniziative di ricerca tendono a suscitare più curiosità conoscitiva di quanta ne soddisfino, come viene esplicitamente dichiarato nella premessa allo studio sulla demografia industriale: "Tutti i dati ottenuti hanno per il momento un valore storico, cioè riferito ad un particolare intervallo di tempo (1977/82). Tuttavia l'obiettivo finale è più ambizioso: se fosse possibile, dopo le opportune messe a punto, trasformare una rilevazione fatta "una tantum" in qualcosa di sistematico e ricorrente si potrebbero fornire, in tempi rapidi, le indicazioni sui mutamenti più importanti che investono con continuità il tessuto industriale piemontese".

2.3.1 La demografia industriale

La prima delle ricerche cui si è fatto riferimento, quella sulla demografia industriale del Piemonte, si propone di definire il grado di "mobilità" dell'apparato produttivo attraverso la quantificazione dell'andamento della "nascita" di nuove imprese e dell'incidenza del corrispondente grado di "mortalità" ed esaminando i risultati in rapporto al settore di attività, alle classi di dimensione occupazionale delle imprese, alle aree di localizzazione delle stesse. Le dinamiche occupazionali e i relativi flussi corrispondenti alla natalità e mortalità delle imprese rappresentano, in tale impostazione, un risultato non certo secondario, ma di secondo grado, ovvero derivato rispetto alla finalità primaria di evidenziare il grado di vivacità imprenditoriale che agisce nel sistema produttivo.

Il periodo cui l'indagine fa riferimento (1977-1982) si pone a cavallo rispetto all'avvio (ottobre 1980) della grande ristrutturazione dell'apparato produttivo industriale e consente di coglierne il grado di "mobilità" con riferimento a due scenari distintamente caratterizzati: il primo, quello iniziale, contraddistinto dalla tendenza al decentramento produttivo e dal forte sviluppo della piccola e media dimensione d'impresa; il secondo, caratterizzato da rilevanti contrazioni occupazionali nelle grandi dimensioni d'impresa e da una forte dinamica delle unità produttive di minori dimensioni con effetto complessivo di redistribuzione dei posti di lavoro disponibili su un ventaglio più ampio di imprese e di riassetto dell'organizzazione produttiva a livello dell'intero apparato manifatturiero.

Rispetto alle finalità conoscitive della ricerca in esame e tenendo conto delle ricordate ulteriori esigenze di poter disporre di informazioni sistematiche e continue in merito, si può ritenere che una risposta complessiva ed esauriente, sia pure a un livello di minore disaggregazione del dato, possa essere riconosciuta nella iniziativa del CERVED con la pubblicazione di Movimprese che fa riferimento, peraltro, ai dati del Registro Ditte delle C.C.I.A.A.. Il minor dettaglio dell'informazione è particolarmente riscontrabile, nel

confronto a livello di distribuzione per classi dimensionali d'impresa, assente nella pubblicazione CERVED e che rappresenta invece, nella ricerca sulla demografia industriale, il più consistente approfondimento dell'elaborazione in termini di flusso con la costruzione delle matrici di transizione per classi di addetti.

Del resto è praticamente inevitabile che le elaborazioni sistematiche e continue presentino un minor grado di dettaglio, in particolare nella fase di avvio, e la iniziativa CERVED è tuttora in sviluppo, presenta un elevato livello di dettaglio di classificazione delle attività economiche ed è estesa pressoché all'intero territorio nazionale.

2.3.2 I flussi occupazionali e localizzativi dell'industria manifatturiera

La seconda ricerca citata come esempio di sperimentazione su dati amministrativi in Piemonte è quella relativa ai flussi occupazionali e localizzativi dell'industria manifatturiera piemontese.

La finalità di tale analisi è duplice: evidenziare le modalità e i percorsi che caratterizzano la forte flessione occupazionale che si è manifestata in Piemonte, nell'industria manifatturiera, nel quinquennio 1980-1985 durante la fase di rapida ed intensa ristrutturazione dell'apparato produttivo; costruire un archivio delle unità locali, a partire dai 10 addetti, con precisi e dettagliati riferimenti di localizzazione territoriale che consenta di determinare consistenza e dinamica dell'apparato produttivo ad ogni livello di zonizzazione definibile, a partire dal modulo comunale elementare.

Date tali finalità l'impostazione metodologica non ha potuto essere orientata a costruire un archivio di dati agevolmente aggiornabili e quindi suscettibile di utilizzo per ottenere elaborazioni sistematiche e continue. Ciò in quanto i dati di fonte amministrativa sono riferiti alle imprese e alla localizzazione della sede legale che, nel caso dei gruppi o delle imprese multilocalizzate, non sempre e non in dettaglio coincide con la localizzazione delle

attività produttive, ovvero delle unità locali, scelte come riferimento di tutte le informazioni disponibili.

I dati di flusso ricavati si riferiscono essenzialmente alle dinamiche occupazionali e delle unità locali: tali informazioni consentono di scomporre le dinamiche complessive nelle componenti di riduzione ed aumento, a loro volta suddivisi nei diversi elementi determinanti e riferiti a ciascuna classe d'ampiezza occupazionale delle unità produttive.

Rispetto alla ricerca sulla demografia industriale prima esaminata, si ha qui una inversione nell'ordine delle finalità conoscitive: i flussi occupazionali, qui primari sono là derivati; i flussi delle unità produttive là primari, sono qui derivati.

Ma, mentre l'esigenza e la finalità principale che ha motivato l'approfondimento di analisi sulla demografia industriale trova risposte continue e sistematiche in MOVIMPRESE, la domanda e l'esigenza di disporre di informazioni organizzate in termini di flusso a scala regionale, che si origina in base alla sperimentata validità dei risultati sia pur parziali e di settore determinati nell'ambito della ricerca sui flussi occupazionali e localizzativi, non trova alcun riscontro nella produzione di statistiche correnti.

3. CONCLUSIONI: I PROBLEMI E LA FATTIBILITA'

La costruzione della matrice dei flussi delle forze di lavoro per la Regione Piemonte, a partire dalle informazioni raccolte con la rilevazione trimestrale ISTAT, è un obiettivo motivato da esigenze di integrazione e di affinamento degli strumenti di analisi delle dinamiche occupazionali e del mercato del lavoro attualmente disponibili e non sempre sufficienti al fine di cogliere appieno il senso delle profonde e rapide trasformazioni in atto.

I problemi che possono costituire ostacolo, sul piano tecnico-organizzativo, alla realizzazione di tale progetto possono essere individuati in relazione a:

- a) reperibilità e disponibilità delle informazioni di base memorizzate su supporto magnetico: per quanto riguarda questo punto è stata accertata la disponibilità dell'ISTAT a fornire tali informazioni nelle forme ritenute adatte allo scopo; il destinatario istituzionale di tali informazioni viene individuato dall'ISTAT nell'ORML della Regione Piemonte che, da parte sua, manifesta interesse per iniziative che si propongono di realizzare gli obiettivi qui indicati.
- b) acquisizione dei programmi di elaborazione dei dati utilizzati dall'ISTAT e definiti in base alla metodologia adottata a scala nazionale: è stato appurato che il Servizio Statistica della Regione Lombardia dispone di tali programmi che ha utilizzato per le prime sperimentazioni a scala regionale (1980) e continua ad utilizzare per le realizzazioni tuttora in corso.
- c) necessarie modifiche dei programmi di elaborazione dati per renderli compatibili con la scala regionale di nuovo riferimento e con il sistema di funzionamento dell'elaboratore di cui si dispone, nel caso sia diverso da quello utilizzato dall'ISTAT a scala nazionale: si tratta di problemi che possono essere compiutamente affrontati e risolti soltanto in corso d'opera, ma è possibile ricavare alcune valutazioni in merito sulla base della precedente esperienza della Lombardia: la fase di adattamento dei programmi

alla scala regionale non pare aver presentato, nel contesto Lombardo, problemi di particolare complessità e la questione del tipo di elaborazione disponibile non si è posta per la disponibilità, in Lombardia, di uno strumento analogo a quello utilizzato dall'ISTAT. Quest'ultimo problema, quand'anche si ponesse, non appare comunque di difficoltà insormontabile.

L'operazione di sovracampionamento regionale è stata effettuata in Piemonte fin dalla rilevazione del gennaio 1979 in misura proporzionale a quella effettuata in Lombardia nel 1980: tale operazione è basilare, ai fini qui proposti, per poter garantire ai risultati ottenuti a scala regionale lo stesso grado di significatività rispetto alla matrice dei flussi nazionale.

Dall'insieme delle considerazioni fatte e qui schematicamente riprese emergerebbero dunque margini sufficientemente fondati di fattibilità del progetto proposto.

Poichè la condizione di fattibilità è certamente necessaria, ma altrettanto certamente non sufficiente rispetto a prospettive di concreta realizzazi del progetto stesso, sono state dedicate ampie parti del presente rapporto a sottolineare anche un'altra condizione importante, quella di utilità.

- Accornero A., "Disoccupazione: nel pelago dei luoghi comuni", in Politica ed Economia, Cespe, n.4, aprile 1986.
- Altieri G. e Pugliese E., "Considerazioni sulle attuali dinamiche del mercato del lavoro in Italia", in Economia e Lavoro, Marsilio Editori, anno XX, n.1, pagg.99-110.
- Bruni M. e Franciosi F.B., "Il mercato del lavoro in Italia: una analisi di flusso", in: L'offerta di lavoro in Italia, a cura di Schenkel M., Marsilio Editori, ottobre 1984.
- Brunetta R. e Venturini A., "La variabile temporale nella transizione tra società industriale e post-industriale", in: Economia e Lavoro, anno XX, n.1, Marsilio Editori, gennaio-marzo 1986.
- Becchi Collidà A., "I giovani a Napoli e..l'occupazione", in Politica ed Economia, Cespe, n.4, aprile 1986.
- Contini B., Ascoli I. e Revelli R., "Studio sulla demografia industriale del Piemonte", quaderno 2.86, Federpiemonte, Torino, 1986.
- Carmignani F., "La disoccupazione non è povertà" in Politica ed Economia, Cespe, n.10, ottobre 1986.
- Cerved, "MOVIMPRESE, movimento anagrafico delle imprese italiane", nuova serie, semestrale, n.1, 1985 (premessa: presentazione della nuova serie di MOVIMPRESE, pag.7; nota metodologica, pagg.135-137).
- Ciravegna D., "Un inquadramento teorico sulle variabili rilevanti ai fini dello studio dell'offerta di lavoro", Relazione al Seminario "Mercato del lavoro: realtà statica o dinamica?", Unione Industriale Torino, ottobre 1987.

Draghi S., "Il modello dei flussi per capire il voto", in Politica ed Economia, Cespe, n.9, settembre 1987.

Frey L., "La disoccupazione in Italia", relazione al convegno AIEL in "I caratteri della disoccupazione", atti della seconda riunione scientifica AIEL, Pavia, aprile 1987.

Ires del Piemonte, "Flussi occupazionali e localizzativi dell'industria manifatturiera Piemontese - 1. La Provincia di Torino", Attività di Osservatorio n.3, gennaio 1987.

Ires del Piemonte, "Piemonte '86 - Relazione sullo stato sociale ed economico della regione", Torino, 1986.

Maugeri C., "L'offerta di lavoro nell'Area Metropolitana Torinese", relazione al Seminario omonimo organizzato dall'ORML-Regione Piemonte, Torino, maggio 1987.

Marotta G. e Pugliese E., "L'Italia della mobilità", in Politica ed Economia, Cespe, n.1, gennaio 1986.

Miceli R. e Gramaglia F., "L'offerta di lavoro: da informazioni aggregate a informazioni sugli individui", dattiloscritto, 1981.

Moriani C., "Aspetti evolutivi dell'occupazione nell'indagine sulle forze di lavoro", in Economia e lavoro, Marsilio Editori, anno XX, n.1, gennaio-marzo 1986.

ORML Regione Piemonte, "Andamento e valutazioni della esperienza di mobilità per i lavoratori Fiat in CIGS", Torino, giugno 1983

Pugliese E., "Ancora su disoccupazione e luoghi comuni", in Politica ed Economia, Cespe, n.7-8, luglio-agosto 1986.

- Petsimeris P., "Processi di deconcentrazione urbana nel contesto piemontese", in Appunti di Politica Territoriale, CIPED, Torino, n.1, gennaio 1987.
- Rosti L., "L'analisi dei flussi da e verso la disoccupazione in Italia 1979/80 - 1983/84", in "I caratteri della disoccupazione", atti della seconda riunione scientifica AIEL, Pavia, aprile 1987.
- Rosti L., "L'analisi dei flussi sul mercato del lavoro in Lombardia e in Italia", relazione presentata alla VII Conferenza Italiana di Scienze Regionali, Urbino, settembre 1986.
- Rosti L., "Imprenditori, per scelta ovvero per necessità", in Politica ed Economia, Cespe, n.2, febbraio 1986.
- Regione Lombardia, Rapporto sul mercato del lavoro lombardo 1986, a cura dell'osservatorio sul mercato del lavoro: parte seconda, "Il mercato del lavoro lombardo ed italiano - un confronto sui dati di stock e di flusso", pagg.95-119.
- Sylos Labini P., "Le classi sociali negli anni '80", Laterza, 1986.
- Schadee M.M.A. e Corbetta P., "Metodi e modelli di analisi dei dati elettorali", ricerche e studi dell'Istituto Cattaneo, il Mulino, 1984.
- Viano F., "Disoccupati a Torino", in Sisifo, n.11, Istituto Gramsci piemontese, Torino, settembre 1987.

Appendice 1

La procedura ISTAT (Moriani, 1981).

tratto da: ISTAT, Supplemento al Bollettino mensile di Statistica, anno 1981, n° 15.

1. PREMESSA

L'Istituto Centrale di Statistica ha elaborato a partire dal 1971 stime dei movimenti della popolazione da una condizione all'altra, osservabili mediante il collegamento di informazioni raccolte sistematicamente con le rilevazioni trimestrali delle forze di lavoro (1).

La pubblicazione di tali dati fu interrotta nel 1977 a seguito di profonde innovazioni nei concetti, nelle definizioni, nelle classificazioni e nelle procedure utilizzate per una serie di rilevazioni sulle forze di lavoro.

Il consolidamento dei risultati ottenuti con la riorganizzazione dell'indagine di base ha permesso di sviluppare le ricerche sui flussi di popolazione inerenti al mondo del lavoro, e quindi di elaborare nuove e migliori procedure per la determinazione dei cambiamenti di condizione osservabili negli individui nel corso del tempo (un trimestre od un anno) mediante l'analisi dei dati raccolti con le rilevazioni trimestrali.

La presente pubblicazione riporta sotto forma di tabelle di serie storiche e di matrici i principali risultati delle elaborazioni fatte, riguardanti il periodo gennaio 1979-gennaio 1981. Matrici più analitiche, sul modello della tavola 10 riportata in appendice, sono a disposizione dei ricercatori che volessero approfondire lo studio dei flussi e potranno essere fornite dietro semplice richiesta e rimborso delle spese vive all'Istituto. L'aggiornamento delle stime sarà fatto trimestralmente ed i relativi dati saranno inseriti nel fascicolo trimestrale «Rilevazione delle forze di lavoro», della collana dei supplementi al Bollettino mensile di statistica,

Nota redatta dal dr. Claudio Moriani del Servizio delle statistiche del lavoro e dei bilanci di famiglia.

(1) ISTAT, *Forze di lavoro e flussi di popolazione*, Supplemento al Bollettino mensile di statistica, n. 5, maggio 1974 (nota a cura della dr.ssa Grazia Giacomini Arangio-Ruiz).

a partire da quello relativo all'indagine del gennaio 1981.

Ulteriori elaborazioni sono allo studio e saranno attuate nel più breve tempo possibile. Le principali riguardano la costruzione di tavole dei flussi per le regioni che hanno un'ampiezza demografica elevata, uno spaccato per età delle tavole nazionali, allo scopo soprattutto di mettere in luce i flussi riguardanti i giovani, ed infine la ricerca dei flussi riguardanti alcune particolari categorie di persone (gli occupati che cercano un'altra occupazione, le non forze di lavoro disposte a lavorare a particolari condizioni, ecc.).

2. NATURA E FINALITÀ DELLA RICERCA

Le forze di lavoro sono costituite, come è noto, dalle persone in età di 14 anni ed oltre che hanno un'occupazione o la cercano attivamente. Esse, al pari della popolazione di cui sono una parziale e particolare specificazione, costituiscono un fenomeno a carattere dinamico: anche se le indagini puntuali ne danno, per forza di cose, una rappresentazione statica, il loro ammontare e la loro struttura interna sono soggetti a continue modificazioni per il mutare della condizione degli individui che fanno parte della popolazione. L'appartenenza alle forze di lavoro, lo stato di occupazione e lo stato di disoccupazione non sono dati immutabili. Nella realtà, infatti, si verifica un continuo interscambio di situazioni che si realizza nel duplice e contrapposto flusso di ingresso nel mondo del lavoro di quanti ne erano in precedenza al di fuori (persone in età non lavorativa, studenti, casalinghe, ecc.) e flusso di uscita (per morte, pensionamento, ecc.) di quanti già ne facevano parte.

Il complesso delle forze di lavoro risulta così rinnovato incessantemente. Ma un continuo cambiamento di situazioni avviene anche all'interno di questo aggregato sociale a seguito del variare del

tipo di attività lavorativa esercitata da una stessa persona, oppure per la perdita dell'occupazione da parte di persone già occupate o per il conseguimento di un'occupazione da parte di persone non occupate. Risulta così rinnovata di continuo anche la composizione qualitativa delle forze di lavoro.

Da quanto detto discende che le variazioni da un trimestre all'altro, o da un anno all'altro, calcolate come differenze tra i dati di consistenza delle forze di lavoro — nelle loro varie specificazioni — rilevati con le indagini campionarie in due istanti diversi, non sono sufficienti ad illuminare compiutamente la dinamica che caratterizza il mondo del lavoro. Esse sono infatti soltanto saldi di flussi contrapposti, che possono assumere dimensioni e qualificazioni molto diverse pur dando nel loro congruaglio lo stesso risultato numerico. In particolare i saldi espressi dalle variazioni nette degli stocks non informano sulla provenienza delle persone che si muovono nella sfera del lavoro e sulla destinazione dei flussi di uscita. Le informazioni di questo tipo, che forniscono essenzialmente dati sui cambiamenti nello stato socio-economico degli individui, possono essere ottenute soltanto rilevando i flussi di popolazione con riferimento al mondo del lavoro.

Gli scopi della ricerca sui flussi non sono soltanto conoscitivi: la loro determinazione apre infatti la strada all'uso di modelli matematici interpretativi della dinamica del mercato del lavoro e di modelli previsivi.

L'interesse alla conoscenza dei flussi si accresce per l'Istat che organizza le rilevazioni delle forze di lavoro, e indirettamente anche per gli utilizzatori dei relativi dati, in quanto la loro ricerca permette di verificare l'efficienza delle stesse rilevazioni, con particolare riguardo al problema delle cadute delle famiglie nel campione ed agli errori di risposta da parte degli intervistati.

3. METODOLOGIA DI CALCOLO

Come nella precedente esperienza richiamata nella Premessa, anche nella presente ricerca l'Istat ha prescelto per la determinazione dei flussi di popolazione nel mondo del lavoro il metodo del confronto delle risposte fornite da uno stesso campione di individui in epoche diverse. Da tale confronto è possibile dedurre i cambiamenti di situazioni avvenute nel periodo trascorso tra i due momenti di ri-

levazione e quindi quantificare i passaggi da una condizione ad un'altra.

Il metodo seguito sfrutta il criterio delle indagini longitudinali, consistente nell'intervistare più volte nel tempo le stesse persone: e ciò è possibile per il nostro Paese grazie allo schema di rotazione delle famiglie nel campione dell'indagine sulle forze di lavoro, come si spiegherà meglio più avanti.

Esistono metodologie alternative che non è stato possibile estendere sistematicamente al materiale informativo raccolto trimestralmente con le rilevazioni sulle forze di lavoro. Tali metodologie si fondano su domande retrospettive poste agli intervistati di una qualsiasi indagine. Con tali domande si accerta la condizione degli individui ad una data precedente (un anno prima, cinque anni prima, ecc.) e si ricreano così le possibilità di confronto tra due situazioni riguardanti gli stessi individui, offerte come si è visto dalle indagini longitudinali. Esempi di applicazione di tale criterio — talvolta non sfruttato per la determinazione dei flussi — si ritrovano nella stessa rilevazione campionaria delle forze di lavoro, nella edizione armonizzata al livello della CEE che si effettua con uno speciale questionario ogni due anni nel mese di aprile (1), e nei censimenti della popolazione (2).

Le domande retrospettive possono riguardare non solo lo status degli individui ad un dato istante del passato ma anche i cambiamenti di status verificatisi per le persone intervistate in un dato periodo di tempo. Questa metodologia viene applicata nell'indagine retrospettiva sulle esperienze di lavoro della popolazione americana, accoppiata nel mese di marzo alla consueta indagine sulle forze di lavoro eseguita negli Stati Uniti; ed in questa stessa indagine, nel caso particolare delle forze di lavoro che hanno cessato di lavorare negli ultimi 12 mesi e negli anni precedenti.

Le metodologie fondate su domande retrospettive, oltre che appesantire con altri quesiti i questionari di rilevazione, si espongono agli errori creati dai ricordi imperfetti degli intervistati, tanto più gravi quanto più remoto nel tempo è il momento di riferimento dei dati o quanto più lungo è il periodo di osservazione dei cambiamenti di status.

(1) Per l'analisi di tali informazioni cfr. M. MARTINI, «La mobilità del lavoro. Un'analisi comparativa. Lombardia, Italia e alcuni Paesi europei», IRER, Milano, 1978.

(2) ISTAT, «11° Censimento generale della popolazione - 24 ottobre 1971: Risultati degli spogli campionari, altri caratteri», Vol. IX, Tomo 2.

Ma anche la metodologia prescelta dall'Istat nella presente ricerca non è esente da inconvenienti. E' universalmente riconosciuto (1) che i dati dei flussi ottenuti mediante accoppiamento di risposte fornite da uno stesso gruppo di individui sono soggetti a tre tipi fondamentali di distorsioni. Il primo origina dal fatto che il sub-campione di individui compresenti in due indagini a confronto, essendo molto più piccolo del campione originario della normale rilevazione delle forze di lavoro, ha un errore campionario diverso da questo e certamente più elevato. Le conseguenze che immediatamente ne derivano sono che il saldo tra il flusso di entrata ed il flusso di uscita diverge in cifra assoluta dalla differenza tra i dati di consistenza, a meno di non usare particolari tecniche di quadratura; e inoltre, essendo comprese tra gli individui esclusi dall'accoppiamento persone non rintracciate ad una delle due date di indagine (per trasferimento di domicilio, per assenza, per rifiuto, ecc.), è molto probabile che le stime campionarie ottenute da un campione più piccolo siano distorte, a meno che gli esclusi non abbiano lo stesso comportamento medio rispetto al lavoro che hanno gli individui accoppiati.

La seconda causa di distorsione è costituita dagli errori di classificazione che gli intervistati compiono nelle interviste successive. Se un individuo si dichiara in cerca di lavoro nella rilevazione, poniamo, di gennaio 1981, e poi assicura di essere fuori delle forze di lavoro nell'indagine successiva di aprile dello stesso anno, pur essendo ancora nell'attesa di un posto promesso ma che non arriva ancora, l'accoppiamento segnerà un movimento di uscita delle forze di lavoro che è sicuramente errato. Ora, che gli intervistati — o i rilevatori che li interrogano — commettano non rari errori di classificazione è un fatto emerso nella esperienza di tutti gli istituti di statistica che eseguono le rilevazioni delle forze di lavoro. L'incidenza di tali errori è aggravata dal fatto che per forza di cose l'intervista non si realizza con tutti i membri della famiglia-campione, ma di solito con il membro che si trova in casa al momento della visita dell'intervistatore, o risponde al telefono (le interviste telefoniche costituiscono negli Stati Uniti oltre il 60% del totale). Ma mentre nei dati di stocks tali errori

tendono a compensarsi, in quelli di flusso tendono invece a cumularsi (2).

La terza causa di distorsione è verificata nelle statistiche degli Stati Uniti ma non ancora accertata in quelle italiane, ed è legata da un condizionamento che riceverebbero le persone intervistate ripetutamente. Ad esempio gli americani trovano che il tasso di disoccupazione delle persone intervistate per la prima volta è più alto di circa il 10% di quello medio di tutte le altre persone che nella stessa indagine sono intervistate per una volta successiva.

Nonostante questi inconvenienti la quantificazione dei flussi derivata dalle rilevazioni delle forze di lavoro sta entrando nella prassi normale degli istituti nazionali di statistica e degli altri enti governativi che la eseguono, come dimostrano le esperienze degli Stati Uniti, del Canada e della Finlandia, che adottano metodologie analoghe a quella seguita nella presente ricerca.

4. PROCEDURA DI ELABORAZIONE

L'abbinamento mediante elaboratore di records individuali relativi ad un medesimo contingente di persone, rilevato in due indagini successive (flussi trimestrali) o corrispondenti, vale a dire condotte nello stesso mese di due anni successivi (flussi annuali), è reso possibile dal sistema di rotazione delle famiglie-campione introdotto nella rilevazione delle forze di lavoro nel gennaio 1971.

4.1 La rotazione delle famiglie

Le famiglie appartenenti al campione delle forze di lavoro sono estratte a giugno di ogni anno dalle anagrafi comunali e sono contemporaneamente suddivise in nove gruppi, più uno di riserva.

Questi nove gruppi, detti sezioni, sono quindi utilizzati a quattro a quattro nelle successive inda-

(2) Si riporta un esempio ripreso da NATIONAL COMMISSION ON EMPLOYMENT AND UNEMPLOYMENT STATISTICS, «Counting the Labor Force», op. cit.: si abbiano due intervistati, uno occupato e l'altro disoccupato in due successive indagini. Se alla seconda indagine entrambi commettono un errore di classificazione invertendo i rispettivi ruoli, le stime degli stocks non ne soffriranno perché gli errori si compenseranno e si continuerà a registrare alla due date lo stesso numero di occupati e di disoccupati. Al contrario le stime dei flussi registreranno due movimenti di fatto mai avvenuti, come cambiamenti da occupato a disoccupato e da disoccupato ad occupato.

(1) NATIONAL COMMISSION ON EMPLOYMENT AND UNEMPLOYMENT STATISTICS, «Counting the Labor Force», Washington, 1979.

gini trimestrali secondo uno schema che consente di avere in comune metà del campione di famiglie tra due indagini consecutive e tra due indagini corrispondenti da un anno all'altro.

Lo schema di rotazione dei gruppi (indicati con lettere latine) è il seguente:

Mesi dell'indagine		Sezioni interessate all'indagine			
		1°	2°	3°	4°
Luglio	1980	S	Z	T	A
Ottobre	1980	T	A	U	B
Gennaio	1981	U	B	V	C
Aprile	1981	V	C	Z	D
Luglio	1981	Z	D	A	E

Come si può vedere, la terza e quarta sezione di un'indagine diventano rispettivamente la prima e la seconda della successiva, mentre per il resto i due campioni sono differenti tra loro. Analogamente, la seconda e quarta sezione di un'indagine diventano rispettivamente la prima e la terza dell'indagine corrispondente un anno dopo, mentre i restanti individui componenti il campione sono diversi.

Ogni anno, inoltre, in occasione dell'indagine di luglio, vengono sostituiti circa un terzo dei comuni con popolazione inferiore a 20.000 abitanti; questo fatto riduce la frazione di famiglie accoppiabili sia nel confronto tra le due rilevazioni successive di aprile e luglio sia in quelli relativi alle rilevazioni ricadenti nello stesso mese di due anni consecutivi. Le famiglie intervistate nella quota di Comuni che restano nell'indagine annualmente rappresentano circa il 18% del campione complessivo e quindi la frazione accoppiabile di famiglie nei casi sopraindicati scende dal 50% al 41% circa.

Nella stessa ottica è necessario tenere conto degli ampliamenti campionari effettuati in numerose provincie al fine di ottenere dati rappresentativi a livello provinciale successivamente al gennaio 1977 data di inizio della nuova serie di indagini (1). Tali ampliamenti, infatti, portano a difficoltà di ordine

(1) Dal gennaio 1977 al gennaio 1981 si sono verificati i seguenti ampliamenti del campione:

- a) luglio 1978 Molise (solo fino a ottobre 1978) e provincia di Torino.
- b) gennaio 1979 Piemonte.
- c) gennaio 1980 Lombardia, provincie di Bologna e Piacenza, Liguria (solo fino ad ottobre 1980).
- d) gennaio 1981 Friuli, Emilia-Romagna, e provincie di Foggia e Brindisi.

pratico in fase di accoppiamento dei records individuali soprattutto per quanto concerne l'interpretazione dei codici indicativi dei Comuni e delle famiglie. Inoltre è intuitivo che nei primi tempi l'elaborazione dei flussi non beneficia dell'ampliamento del campione in quanto per forza di cose l'accoppiamento avviene solo per le parti in comune alle due indagini a confronto.

4.2 La numerazione delle famiglie

I questionari relativi alle famiglie intervistate sono numerati con due sistemi:

1) una numerazione progressiva nell'ambito di ciascun Comune;

2) una numerazione progressiva nell'ambito di ciascuna sezione per area di rilevazione di uno stesso Comune.

Per ridurre l'influenza di errori di numerazione in sede di compilazione dei questionari da parte dei rilevatori o in sede di trasferimento dei dati su supporto magnetico ai fini dell'accoppiamento delle informazioni relative a ciascun individuo ci si è basati sul secondo sistema, in quanto per il primo (cioè numerazione progressiva nell'ambito di ciascun Comune) si è constatato che è sufficiente uno slineamento di un solo codice per rendere inutilizzabili ai fini dell'elaborazione tutte le informazioni di un Comune.

Inoltre, con il secondo sistema, è sufficiente rifarsi alla sola indicazione della sezione per mettere a confronto in blocco tutti gli individui che la compongono.

4.3 La procedura di accoppiamento

Per giungere all'associazione degli individui si mettono dapprima in corrispondenza i Comuni, previa eliminazione di quelli che hanno partecipato ad una sola delle due indagini messe a confronto.

Quindi basandosi sulla numerazione per area di rilevazione (o di sezione, se il Comune, essendo piccolo, forma una sola area di rilevazione), si procede all'accoppiamento per elaborazione delle famiglie, e nell'ambito di ciascuna, all'associazione dei singoli individui sulla base delle rispettive età, con la necessaria tolleranza per tenere conto del tempo intercorrente tra un'indagine e l'altra.

Le principali informazioni relative a quest'operazione sono indicate in un tabulato sul quale figurano per ciascuna delle due indagini a confronto i dati comunali per sesso relativi a:

- 1) popolazione;
- 2) individui scartati perché appartenenti a sezioni non comuni alle due indagini;
- 3) individui scartati perché l'intera famiglia non aveva riscontro nell'altra indagine;
- 4) individui non accoppiati nell'ambito di una stessa famiglia, di cui nati nel periodo;
- 5) individui accoppiati.

Una volta accertata l'esattezza delle informazioni ottenute si passa ad operare sul gruppo degli individui accoppiati che, essendo un sub campione, necessita di essere riportato al proprio universo di riferimento.

4.4 L'universo di riferimento degli individui accoppiati

L'universo di riferimento degli individui accoppiati diverge dalla popolazione complessiva cui si riferisce una qualsiasi rilevazione delle forze di lavoro (1). Esso infatti è costituito dalla popolazione residente (compresi gli individui temporaneamente all'estero) corrispondente alla seguente espressione:

$$P_a = P_1 - N - I \quad [1]$$

dove

P_a = popolazione residente accoppiata

P_1 = popolazione residente complessiva al tempo 1

N = nati

I = immigrati dall'estero e iscritti per cambio di residenza.

(1) La popolazione di riferimento per l'indagine sulle forze di lavoro è la popolazione residente come risulta dalle anagrafi comunali, nel senso che popolazione al tempo 1 è uguale alla popolazione al tempo 0 più i nati vivi, meno i morti, più gli iscritti da Comuni appartenenti a strati diversi o dall'estero meno i cancellati per Comuni appartenenti a strati diversi o per l'estero.

Tale popolazione, in quanto popolazione residente, comprende anche i temporaneamente assenti, che entrano nel calcolo per la determinazione dei coefficienti di espansione. Il dato di popolazione che compare, però, nelle tavole finali, non comprende gli individui temporaneamente assenti che vengono esclusi dalle elaborazioni successive al riporto all'universo. La popolazione che risulta è quindi una popolazione residente presente costituita dal complesso degli individui residenti, rilevati come presenti nel territorio nazionale.

Tenuto conto della nota equazione della dinamica della popolazione

$$P_1 = P_0 + N - M + I - E \quad [2]$$

si ha anche

$$P_1 - N - I = P_0 - M - E = P_a \quad [3]$$

dove i termini nuovi sono

P_0 = popolazione residente complessiva al tempo 0

M = morti

E = emigrati all'estero e cancellati per cambio di residenza.

Si deduce dalla [3] che P_a può essere determinato in due modi diversi ma convergenti. Nell'un caso o nell'altro la relazione [3] si può intendere a due livelli diversi: a) a livello di popolazione nazionale; b) a livello di popolazione dei singoli strati. I nati e i morti saranno gli stessi nei due casi, mentre i movimenti migratori riguarderanno solo quelli con l'estero nel caso a) e saranno comprensivi di quelli interni nel caso b). La procedura adottata, dopo alcuni esperimenti, si è basata sulla relazione [1] intesa a livello nazionale, al quale i movimenti migratori interni restano compensati. In pratica l'universo della popolazione accoppiata è stato determinato sottraendo dalla popolazione più recente delle due considerate i dati riguardanti i nati e gli immigrati dall'estero. Così facendo si è tollerata la distorsione introdotta nei dati di popolazione di ogni singolo strato col trascurare gli effetti dei movimenti migratori interni. Questi tuttavia rappresentano in media, in ogni trimestre, soltanto il 5 per mille della popolazione: e se anche è verosimile che le caratteristiche socio-economiche degli individui che cambiano comune di residenza siano diverse da quelle della restante popolazione, l'irrelevanza del loro peso relativo assicura che i risultati complessivi delle elaborazioni non siano sensibilmente intaccate dalla loro esclusione dal computo esplicito. In pratica il procedimento seguito utilizza il campione degli accoppiati per rappresentare nei singoli strati anche il gruppo di coloro che hanno cambiato residenza (2).

(2) Si è accennato ad alcuni esperimenti effettuati, miranti a stimare gli individui immigrati nei singoli strati. Preimpresso che l'estrazione del campione di famiglie dall'anagrafe si effettua una volta all'anno e che lo schema di rotazione richiama nell'indagine sezioni di famiglie

Ritornando alla relazione [1] e quindi alla determinazione della popolazione accoppiata nei singoli strati, resta il problema della quantificazione dei nati e dei soli immigrati dall'estero. Entrambi questi flussi sono ottenibili dalle statistiche demografiche correnti, direttamente o mediante proiezioni dei dati dei precedenti periodi per le stime riguardanti i trimestri più recenti. Una volta accertati, i nati e gli immigrati sono stati eliminati dalla popolazione al tempo 1 di ciascuno degli strati.

4.5 La matrice di transizione

Dopo essere stati riportati all'universo gli individui accoppiati sono classificati secondo i caratteri socio-economici posseduti al tempo 0 e secondo quelli posseduti al tempo 1, formando la parte centrale di una matrice di transizione. La matrice presenta lungo le righe la ripartizione degli individui secondo i caratteri del tempo 0 e lungo le colonne l'analoga ripartizione degli stessi individui secondo i caratteri del tempo 1. Si ottiene in tal modo la classificazione bivariata secondo l'origine e la destinazione degli individui. Tutte le caselle di tale matrice, diverse da quelle poste sulla diagonale, indicano cambiamenti di condizione socio-economica e contrassegnano quindi fenomeni riguardanti il mondo del lavoro. La matrice degli accoppiati, che sono individui compresenti alle due date di riferimento degli stocks, non è la matrice completa. Questa ha bisogno, per essere perfezionata e « chiusa », di due righe aggiuntive che costituiscono la cornice della tavola. Le colonne riportano i morti e gli emigrati

estratte nel ciclo annuale precedente, sono limitate a priori le possibilità che le rilevazioni delle forze di lavoro registrino compiutamente gli spostamenti di intere famiglie da un Comune all'altro. Tuttavia è teoricamente immaginabile che un certo numero di famiglie trasferite rientri nel campione intervistato nella seconda delle due indagini confrontate nella stima dei flussi, e quindi sia possibile isolarle e riferire le loro caratteristiche socio-economiche al complesso degli immigrati nei singoli strati. Infatti sono stati conteggiati gli individui presenti al tempo 1 ma assenti al tempo 0 per cause diverse (esclusi però i nati nel periodo), e sono stati riportati all'universo. Senonché, nel numero così stimato si sommano agli individui effettivamente immigrati quelli non presenti o non identificabili nella prima indagine per errori di codifica o di numerazione o per cadute o rifiuti di intervista, e pertanto il loro numero ricavato dal campione risulta troppo elevato (circa due milioni di persone) e non può essere utilizzato senza essere sottoposto a nuovi filtri di difficile concezione.

per l'estero del periodo; le righe, i nati e gli immigrati dall'estero. Tali dati sono tutti desumibili dalle statistiche demografiche. Occorre tuttavia ripartire i morti, gli emigrati e gli immigrati secondo i caratteri socio-economici previsti nella tavola. Per questo si è seguita una strada diversa per i morti e per i movimenti migratori. Per i morti si è calcolato prima il numero teorico pertinente ad ogni casella della colonna, applicando i coefficienti di mortalità specifica per sesso e condizione professionale, quindi si è riproporzionato il numero esatto dei morti a quello teorico. Quanto agli emigrati e agli immigrati dall'estero, la loro ripartizione per condizione socio-economica è stata ipotizzata simile a quella delle persone temporaneamente all'estero rispettivamente al tempo 0 e al tempo 1, come risulta dalla stessa rilevazione campionaria. I nati infine sono tutti classificati nella casella dei minori di 14 anni.

Al totale di ciascuna riga della matrice, comprese le uscite per morte ed emigrazione, deve corrispondere il valore della modalità quale risulta dall'indagine delle forze di lavoro eseguita al tempo 0. Allo stesso modo, il totale di ciascuna colonna, compresi gli ingressi per nascita ed immigrazione, deve corrispondere al valore della modalità rilevata dall'indagine al tempo 1. In realtà, la prima compilazione della tabella completa mostra numerose divergenze, che si spiegano con il fatto che la matrice degli accoppiati deriva dal riporto all'universo di un sub-campione di individui (diverso quindi dai campioni completi intervistati alle due date considerate), e la cornice della tavola riporta informazioni esterne all'indagine delle forze di lavoro. La convergenza viene ottenuta facendo ricorso al procedimento utilizzato per la quadratura delle tavole input-output relative ai flussi di beni e servizi nell'economia. Si opera nel seguente modo:

- 1) si inseriscono i risultati relativi alle due indagini delle forze di lavoro messe a confronto;
- 2) si inseriscono i dati relativi agli ingressi (nati e immigrati dall'estero) ed alle uscite (morti, emigrati all'estero), ripartiti per le varie modalità secondo i criteri di cui sopra;
- 3) si ottengono quindi i totali imposti di riga e di colonna della matrice interna degli accoppiati sottraendo le somme delle uscite e degli ingressi dai risultati della rilevazione rispettivamente più lontana o meno lontana nel tempo.

Le differenze che sussistono tra i totali effettivi della matrice di flusso ed i totali imposti così calcolati sono quindi annullate facendo uso del metodo RAS richiamato (1).

4.6 Riepilogo della procedura

Riassumendo, la procedura si compone delle seguenti fasi principali:

a) accoppiamento dei dati individuali relativi alle due rilevazioni a confronto:

- identificazione dei Comuni rotati (solo nel caso di sostituzione dei Comuni) e accoppiamento dei Comuni nei quali sono state effettuate entrambe le rilevazioni a confronto;

- scarto della parte del campione non in comune e rilevazione dei dati di popolazione necessari al successivo riporto all'universo;

- accoppiamento dei singoli individui sulla base del Comune, della famiglia e dell'età;

- correzione dei dati anomali;

- emissione dei tabulati di controllo.

(1) Questo metodo, di comune applicazione, consiste nel riproporzionamento degli elementi di una matrice quadrata in modo da eguagliare le somme di riga di A,

$$\sum_{j=1}^n a_{ij}, i = 1, 2, \dots, n$$

e le somme di colonna di A,

$$\sum_{i=1}^n a_{ij}, j = 1, 2, \dots, n$$

a dei valori forniti in precedenza.

Ciò si ottiene moltiplicando ciascun elemento di A, indicato con a_{ij} , per un coefficiente di riga r_i , e per un coefficiente di colonna, s_j . La serie dei $2n$ coefficienti si trova con un metodo iterativo, la cui convergenza ad una soluzione è dimostrata, per matrici non negative, con la sola ovvia condizione che la somma delle somme di riga imposta sia uguale alla somma delle somme di colonna imposte.

In pratica si utilizza un metodo iterativo, per aggiustare direttamente gli elementi della matrice A, ed i coefficienti r e s risultano solo indirettamente dall'applicazione di tale procedimento. In effetti, si uguagliano dapprima i totali di riga calcolati sulla matrice originale A ai valori imposti dall'esterno, ottenendo così una matrice $A^{(1)}$ che soddisfi al vincolo sulle righe, moltiplicando

gli elementi di A per $a_{i1} / \sum_{j=1}^n a_{ij}, i = 1, 2, \dots, n$ dove

gli a_{i1} sono i valori imposti. Analogamente si procede per i totali di colonna moltiplicando gli elementi di $A^{(1)}$ per $a_{1j} / \sum_{i=1}^n a_{ij}, j = 1, 2, \dots, n$. Si ottiene così una matrice $A^{(1)}$

che rispetta i vincoli per le colonne, ma probabilmente non rispetta più i vincoli per le righe. Iterando il procedimento con successive quadrature delle righe, delle colonne etc., come si è detto, si giunge ad una fase in cui le differenze tra le somme di righe e colonne così elaborate differiscono da quelle imposte a meno di un valore comunque piccolo.

b) riporto all'universo:

- stampa di un tabulato di controllo con la popolazione per sesso, regione e tipo di Comuni (Comuni capoluoghi, Comuni non capoluoghi con più di 20.000 abitanti, altri Comuni);

- calcolo dei dati di nati e immigrati per la determinazione della popolazione degli accoppiabili;

- calcolo ed inserimento dei nuovi coefficienti di espansione;

- compilazione della matrice degli individui accoppiati.

c) elaborazione delle stime su dati esterni e preparazione delle tavole:

- stima dei nati, morti, iscritti e cancellati e per l'estero;

- compilazione della tavola generale grezza secondo la condizione socio-economica;

- quadratura della tavola con il metodo RAS;

- preparazione delle tavole finali per la stampa in off-set dei valori assoluti e percentuali.

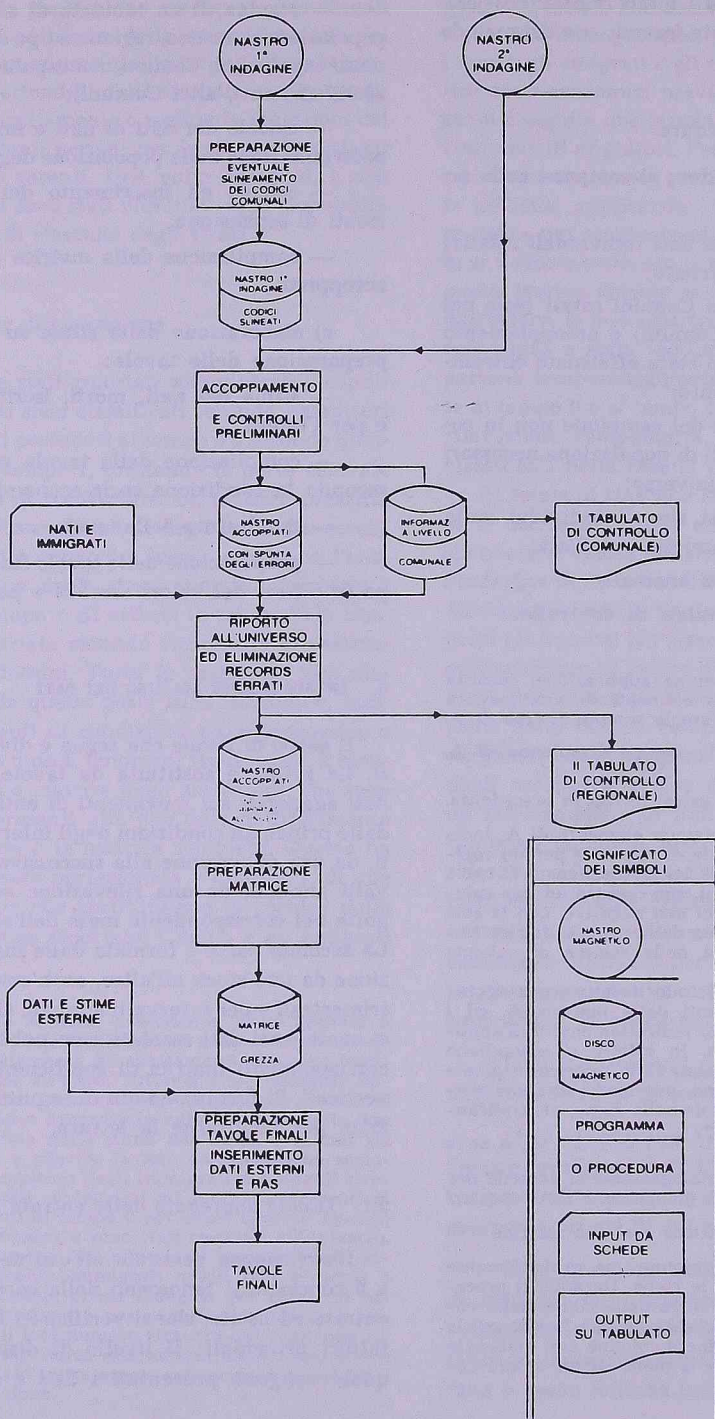
5. INTRODUZIONE ALL'USO DEI DATI

Il corpo di tavole che segue è diviso in due parti. La prima è costituita da tavole che riportano dati aggregati sui movimenti di entrata e di uscita dalle principali condizioni negli intervalli trimestrali, da una rilevazione alla successiva, e negli intervalli annuali, da una rilevazione ad un'altra eseguita nel corrispondente mese dell'anno successivo. La seconda parte è formata dalle matrici di transizione da uno stock all'altro, anch'esse per intervalli trimestrali e per intervalli annuali. Oltre alle matrici contenenti dati assoluti sono pubblicate anche le corrispondenti matrici di coefficienti, orizzontali e verticali. Si forniscono qui di seguito alcune indicazioni per facilitarne la lettura.

5.1 Tavole aggregate delle entrate e delle uscite

L'avvertenza generale all'uso delle tavole da 1 a 5 concerne il fenomeno della compensazione tra entrate ed uscite, che si verifica in funzione di due fattori principali: il livello di disaggregazione al quale vengono presentati i dati e l'ampiezza del-

SCHEMA DELLA PROCEDURA ELETTRONICA



l'intervallo temporale. Circa il primo fattore, maggiore è la disaggregazione più ampi appaiono i volumi dei flussi di entrata e di uscita, e quindi meno incidente è il fenomeno della compensazione. Per un esempio si rimanda alla prima riga della Tav. 1: si può constatare che la cifra degli entrati tra le forze di lavoro nel 1° trimestre del 1979 (1.185.400 unità) è inferiore alla somma degli entrati nelle due categorie che compongono le forze di lavoro, gli occupati e le persone in cerca di occupazione, somma che ammonta a 1.719.500 unità (1.067.400 + 652.100). La differenza tra le due cifre, pari a circa 534.000 persone, è dovuta a movimenti interni tra gli strati di occupazione e di ricerca di lavoro, per cui diventano rilevanti, per determinare il flusso esclusivo di ingresso nelle forze di lavoro, soltanto le entrate provenienti dalle non forze di lavoro.

Per quanto concerne gli effetti di compensazione che crea l'ampiezza dell'intervallo temporale che separa le due rilevazioni a confronto, è intuitivo che più lungo è tale intervallo più frequenti sono i ritorni di individui alla condizione di partenza, in modo che più numerosi sono i casi di compensazione. Per tale motivo la somma dei movimenti tra le varie condizioni relativi ad intervalli trimestrali è superiore alla cifra che esprime lo stesso tipo di movimenti per un periodo annuale che abbracci gli stessi trimestri osservati. Si rimanda ancora una volta alla Tav. 1, prima colonna, per un esempio. La somma degli entrati nello stato di occupazione nei quattro trimestri da gennaio 1979 a gennaio 1980, pari a 4.575.800 lavoratori, è di gran lunga superiore alla cifra che indica le entrate nell'occupazione avvenute nell'intervallo annuale tra gennaio 1979 e gennaio 1980 contrassegnato dalle due rilevazioni invernali corrispondenti. In tale periodo di tempo oltre 2.600.000 persone hanno conseguito e perduto l'occupazione, di modo che soltanto 1.966.200 persone si trovano a distanza di un anno ad aver ottenuto e conservato un posto di lavoro, figurando così nel movimento annuale di ingresso nell'occupazione.

La composizione dei flussi compensati nell'ambito di un certo aggregato si ottiene dall'esame dei dati più dettagliati disponibili nelle tavole riepilogative o direttamente nelle matrici per uno stesso intervallo temporale. La composizione dei flussi che si compensano negli intervalli di tempo annuali è invece deducibile dall'analisi dei dati di flusso rela-

tivi ai trimestri che compongono il periodo annuale studiato.

Ciò premesso è possibile dare qualche indicazione di massima sui principali movimenti osservati.

a) *Entrate ed uscite dalle forze di lavoro.* Tra il 1979 ed il 1980 il complesso delle forze di lavoro italiane è aumentato, in media, di circa 300.000 persone. Questa cifra è la risultante finale di opposti movimenti di entrata e di uscita dalla vita attiva, a seguito dei quali è cresciuta la partecipazione al lavoro della popolazione italiana. Le elaborazioni compiute per quegli anni portano a stimare in circa 2.190.000 gli ingressi nella vita attiva, in parte bilanciati da circa 1.890.000 uscite per morte, emigrazione, pensionamento o semplice ritiro temporaneo.

Per le ragioni che sono state dette prima l'ampiezza reale dei flussi lordi è più elevata perché le cifre indicate misurano i movimenti rappresentati da cambiamenti intervenuti nella consistenza delle forze di lavoro ad un anno di distanza. Questi cambiamenti, a meno dei nati, dei morti e degli immigrati ed emigrati, sono rivelati da rappresentazioni statiche di persone copresenti nelle due date a confronto, distanziate di un anno l'una dall'altra. In media, cioè, 2.190.000 persone, che un anno prima venivano riportate tra le non forze di lavoro, alla seconda rilevazione sono state colte tra le forze di lavoro, e 1.890.000, che erano occupate o cercavano un lavoro un anno prima, sono state colte in seguito tra la popolazione inattiva. Così limitata l'osservazione, non è dato sapere tutti i movimenti di entrata e di uscita delle forze di lavoro che sono accaduti lungo tutto l'arco del periodo annuale osservato e che si sono nel frattempo ricomposti, nel senso che le persone interessate hanno ripreso il vecchio status in entrambe le osservazioni statiche, oppure fanno parte dei due flussi lordi rilevati dopo aver cambiato più di una volta la propria posizione nel mondo del lavoro.

Abbreviando i periodi che intercorrono tra due osservazioni successive è possibile peraltro, come si è già detto all'inizio del presente paragrafo, rendere più completa e in qualche modo più « dinamica » la registrazione dei flussi. La somma dei flussi trimestrali sull'arco di dodici mesi è risultata in media — per quattro intervalli annuali che vanno da gennaio a gennaio, da aprile ad aprile, da luglio a luglio e da ottobre ad ottobre — di oltre 5.200.000 ingressi nelle forze di lavoro e 4.900.000

usciti, con un saldo positivo pari anch'esso a circa 300.000 unità. I movimenti di entrata e di uscita, misurabili per intervalli trimestrali anziché annuali, che risultano alla fine compensati sono dell'ordine di 3 milioni di unità: cifra imponente, essendo pari al 13,5% della consistenza media delle forze di lavoro nel biennio 1979-80.

Ritornando ai flussi lordi calcolati in media sull'intervallo annuale si osserva che essi sono costituiti prevalentemente da donne in entrambe le direzioni. Tra gli entrati, infatti, le donne sono state in media il 56% e tra gli usciti il 54%. La maggiore mobilità della componente femminile è riscontrabile anche nel flusso dei movimenti compensati, dato che su 3 milioni di movimenti quasi 1.900.000 hanno riguardato le donne (63%).

b) *Flussi tra occupazione, ricerca di lavoro e condizioni non professionali.* Tra il 1979 ed il 1980 circa 2.000.000 di persone hanno trovato o ripreso un posto di lavoro, mentre 1.700.000 circa lo hanno perduto o abbandonato. Tra i « nuovi » occupati il 71% circa proviene dalle condizioni non professionali, il 27% dalla ricerca di occupazione e la restante frazione dall'immigrazione. Quanto agli usciti, il 25% di essi era di nuovo alla ricerca dell'occupazione, il 75% circa è rifluito tra le non forze di lavoro e la restante parte non figura più tra gli occupati per morte o emigrazione. I passaggi diretti dalle condizioni non professionali all'occupazione prevalgono nella composizione del flusso di entrata, anche quando si osservano i singoli intervalli trimestrali. Ciò è dovuto in parte al fatto che la transizione per la condizione intermedia della ricerca di lavoro dura obiettivamente meno di un anno (e alle volte meno di tre mesi) per singole iniziative conclusesi positivamente, ed in parte al fatto per molte forme di occupazione (temporanee, stagionali, ecc.) la transizione è breve e immediata.

Rispetto alle consistenze medie del biennio 1979-80 i flussi di entrata nell'occupazione hanno costituito in media il 6,8% per i maschi, il 16,2% per le femmine ed il 9,7% per il complesso dei sessi. A loro volta i flussi di uscita sono stati pari al 6,1% per i maschi, al 13,1% per le femmine ed all'8,3% per il complesso degli occupati.

Per quanto riguarda i movimenti tra i settori di attività economica che avvengono nell'ambito delle forze di lavoro occupate, si rileva un forte in-

terscambio tra l'industria e il terziario e viceversa e nell'insieme un imponente flusso di passaggio da un settore all'altro, che nella media dell'anno si ragguaglia a circa 1.650.000 movimenti. Ripartiti per settore, gli ingressi lordi hanno riguardato per il 16,3% l'agricoltura, il 38,2% l'industria ed il 45,2% le altre attività; le uscite sono state per il 20,5% dall'agricoltura, il 37,3% dall'industria e per il 42,2% dalle altre attività. Questi dati vanno comparati con la composizione della consistenza degli occupati, che nella media dell'anno 1980 è stato: 14,1% agricoltura, 37,6% industria e 48,3% altre attività.

c) *Entrate e uscite dalla ricerca di occupazione.* Il gruppo delle persone in cerca di occupazione è quello che subisce un ricambio più intenso. Su una consistenza media del biennio 1979-1980 di 1.698.000 persone in cerca di lavoro si sono osservati negli intervalli annuali flussi di entrata e di uscita pari a circa 990.000 unità (58%). Gli ingressi nella ricerca nel biennio considerato provenivano in media per il 74% circa dalle non forze di lavoro e per il 26% dall'occupazione; le uscite si erano dirette per il 55% verso la popolazione inattiva e per il 45% verso l'occupazione. Le cifre indicano una frequenza molto alta di rientri tra le non forze di lavoro, che possono essere in gran parte attribuiti a persone che cercano posizioni di lavoro stagionali, non impegnative o di breve durata e che probabilmente le conseguono nell'intervallo di un anno che separa le due rilevazioni degli stocks che si confrontano.

Il tasso di permanenza nella stessa condizione a distanza di un anno è pari al 42%. Si deve tener presente, tuttavia, che tra le persone che si rilevano ancora in cerca di occupazione un anno dopo la prima rilevazione possono essere compresi individui che hanno nel frattempo trovato un lavoro, lavorato per brevi periodi e quindi perduto l'occupazione. I tassi di permanenza sono differenziati a seconda delle singole categorie che compongono il gruppo delle persone in cerca di occupazione, e tendono a crescere nel passaggio dalla categoria delle « altre persone in cerca di lavoro » (23%) ai disoccupati in senso stretto (27% circa), ed infine alle persone in cerca di prima occupazione (45% circa). Le donne presentano in complesso un tasso di permanenza (44,8%) più alto di quello dei maschi (38,4%); nella categoria delle persone in cerca di prima occupa-

zione, peraltro, i due sessi si trovano alla pari, probabilmente perché in tale categoria, caratterizzata da una forte presenza di giovani e di persone fornite di un migliore livello di istruzione, le opportunità di impiego sono pressoché eguali per i maschi e per le femmine. Tassi di gran lunga più elevati rispetto a quelli dei maschi presentano invece le donne nella categoria dei disoccupati in senso stretto ed in quella delle «altre persone in cerca di lavoro».

5.2 Matrici di transizione da una situazione all'altra

Le matrici di transizione da una situazione all'altra, sia per intervalli trimestrali sia per intervalli annuali, sono esposte, oltre che per sesso, sotto forma di cifre assolute, coefficienti orizzontali e coefficienti verticali.

In generale le matrici vanno lette nel seguente modo. Per riga sono riportate per ciascuna modalità indicata nel colonnino i flussi di uscita verso le condizioni indicate nella testata. Il totale di riga costituisce lo stock iniziale, di cui una parte permane nella stessa condizione e si ritrova nella casella della diagonale. Per colonna sono riportati i flussi di entrata in ciascuna condizione indicata nella testata, provenienti da quelle indicate nel colonnino. Il totale di colonna costituisce lo stock finale, di cui una parte si trovava nella stessa condizione fin dall'inizio del periodo di osservazione, come risulta dalla cifra inserita nella casella della diagonale.

Poiché per riga si ha la distribuzione dello stock iniziale secondo le condizioni di cui le persone si rilevano alla fine del periodo osservato, la stessa in caso di permanenza e le altre, diverse da quella iniziale, in caso di cambiamento, le percentuali sul totale di riga esprimono in termini relativi l'entità

dello stock iniziale che non ha cambiato condizione e l'importanza delle uscite nelle loro varie direzioni. I coefficienti orizzontali così calcolati sono interpretabili, perciò, come tassi di uscita dalla situazione iniziale e come tasso di permanenza nella stessa situazione per il coefficiente che si trova nella casella sulla diagonale della matrice, alla congiunzione di riga e colonna relative alla stessa condizione.

I corrispondenti coefficienti verticali calcolabili come percentuali dei dati di colonna sul loro totale sono a loro volta interpretabili come tassi di entrata nella nuova situazione, riferendosi alla struttura dello stock alla fine del periodo secondo la condizione di provenienza. Anche in questo caso il coefficiente che si trova nella casella sulla diagonale è un tasso di permanenza nello stesso stato, riferito questa volta non già allo stock di partenza bensì a quello di arrivo.

In altre parole, il coefficiente orizzontale è assimilabile alla probabilità di passaggio da uno stato di origine ad un qualunque stato di arrivo in una matrice di transizione; mentre il coefficiente verticale ha significato analogo a quello di un coefficiente tecnico di una matrice intersettoriale di beni e servizi e fornisce la percentuale di contributo dato allo stock finale relativo ad una certa modalità fornito dai vari stocks iniziali relativi a tutte le modalità.

AVVERTENZA

Per brevità, nei titoli delle tavole si sono usate le indicazioni: «flussi di entrata e uscita trimestrali», per indicare i flussi tra due indagini successive a distanza di un trimestre, e «flussi di entrata e uscita annuali» per indicare i flussi tra due indagini corrispondenti a distanza di un anno.

Appendice 2

Flussi da elaborazione di stock (evoluzione del modello di Goodman, 1953: uso delle correlazioni fra dati aggregati per stimare le correlazioni a livello individuale).

tratto da: H.M.A. SCHADEE e P. CORBETTA, Metodi e modelli di analisi dei dati elettorali, il Mulino, 1984.

Il modello che viene qui presentato è tale che, se applicato ad un certo tipo di informazioni aggregate, permette un'interpretazione dei coefficienti di regressione in termini di frazioni di popolazione. Tali coefficienti possono essere stimati con le tecniche standard di regressione (criterio dei minimi quadrati), disponibili in un qualsiasi programma di computer che sia in grado di effettuare regressioni con molteplici variabili indipendenti.

Le stime dei coefficienti trovate in questo modo possono anche essere negative o maggiori di 1, fatti questi che non possono essere accettati dal momento che esse sono interpretate come frazioni di popolazione. Se ciò avviene, è perché il modello è stato specificato in maniera incompleta, nel senso che delle variabili essenziali non sono state incluse nelle procedure di stima. Se tali variabili sono disponibili, esse vanno incluse, ma, se non lo sono, le stime dei coefficienti possono essere «forzate» affinché soddisfino ai vincoli imposti al modello. In questo capitolo proporremo una misura, o un insieme di misure, del grado con cui tali stime forzate deviano dalla soluzione ottenuta con le semplici tecniche di regressione non forzate.

Il modello fa riferimento al caso in cui la popolazione di un'unità territoriale viene classificata sulla base di due variabili. La prima variabile, Y , ha K categorie, la seconda, X , ha J categorie. Ogni individuo della popolazione deve appartenere ad una ed una sola delle categorie di Y , e ad una ed una sola delle categorie di X . Sia Y_k la frazione di popolazione dell'unità territoriale che appartiene alla categoria k della variabile Y , e X_j la frazione di popolazione che appartiene alla categoria j della variabile X .

Dal momento che non ci sono frazioni negative di popolazione, e dal momento che una frazione di popolazione non può essere più grande dell'intera popolazione, si deve avere che:

$$[4] \quad 0 \leq Y_k \leq 1 \quad \text{per ogni } k$$

$$[5] \quad 0 \leq X_j \leq 1 \quad \text{per ogni } j$$

Inoltre, dal momento che ogni individuo appartiene ad una categoria sola, la somma di tutte le frazioni di popolazione nelle categorie di una stessa variabile deve essere uguale a 1:

$$[1] \quad \sum_k Y_k = 1 \quad k = 1 \dots K$$

$$[2] \quad \sum_j X_j = 1 \quad j = 1 \dots J$$

Tutti gli individui appartenenti alla categoria k di Y devono appartenere ad una categoria della variabile X . Se noi chiamiamo b_{kj} X_j la frazione di popolazione totale che appartiene alla categoria k della variabile Y ed alla categoria j della variabile X , allora

$$Y_k = b_{k1} X_1 + b_{k2} X_2 + b_{k3} X_3 + \dots + b_{kj} X_j + \dots + b_{kJ} X_J$$

Secondo questo modo di presentare la popolazione, i b_{ki} rappresentano le frazioni della categoria X_i che appartengono alla categoria Y_k . Dal momento che sono frazioni di popolazione, i b_{ki} non possono essere negativi né maggiori di 1:

$$[6] \quad 0 \leq b_{ki} \leq 1 \quad \text{per ogni } k, i$$

e dal momento che ognuno che appartiene alla categoria j di X deve appartenere ad una categoria k di Y :

$$[3] \quad \sum_k b_{ki} = 1 \quad k = 1 \dots K, \text{ per ogni } i$$

Per ogni unità territoriale noi possiamo scrivere il seguente modello per l'intera popolazione:

$$\begin{aligned} Y_1 &= b_{11}X_1 + b_{12}X_2 + \dots + b_{1i}X_i + \dots + b_{1j}X_j \\ Y_2 &= b_{21}X_1 + b_{22}X_2 + \dots + b_{2i}X_i + \dots + b_{2j}X_j \\ &\vdots \\ Y_k &= b_{k1}X_1 + b_{k2}X_2 + \dots + b_{ki}X_i + \dots + b_{kj}X_j \\ &\vdots \\ Y_K &= b_{K1}X_1 + b_{K2}X_2 + \dots + b_{Ki}X_i + \dots + b_{Kj}X_j \\ \hline 1 &= X_1 + X_2 + \dots + X_i + \dots + X_j \end{aligned}$$

Questo modello è logicamente vero per ogni unità territoriale, anche se i coefficienti b_{ki} sono sconosciuti.

Nel caso di più unità territoriali, dove Y_k e X_i sono noti, i b_{ki} possono essere stimati. Per fare ciò tuttavia sono necessari alcuni assunti relativi ai b_{ki} . In primo luogo assumiamo che essi siano esattamente uguali in tutte le unità territoriali. In tal caso un totale di J unità territoriali dovrebbe essere sufficiente per risolvere il sistema rispetto ai coefficienti, poiché le righe del modello forniscono J equazioni nelle J variabili da b_{k1} a b_{kj} , e può quindi ottenersi una soluzione esatta.

Ma è improbabile che i coefficienti b siano esattamente gli stessi in tutte le unità territoriali. Possiamo tuttavia fare un'altra assunzione, e cioè che essi differiscano l'un l'altro di piccole quantità, casualmente distribuite, fra area ed area:

$$[E1] \quad b_{ikj} = g_{kj} + d_{ikj} \quad \text{in ogni } i$$

dove g_{kj} è il valore attorno al quale i coefficienti b_{ikj} , cioè i coefficienti b_{ki} in ognuna delle $i = 1 \dots N$ unità territoriali, variano e d_{ikj} è una piccola componente casuale. Dal momento che g_{kj} rappresenta di nuovo una frazione di popolazione

$$[E3] \quad 0 \leq g_{kj} \leq 1 \quad \text{per ogni } k, j$$

$$[E4] \quad \sum_k g_{ki} = 1 \quad k = 1 \dots K$$

Sotto queste condizioni i g_{kj} possono essere stimati mediante le tecniche standard di regressione (criterio dei minimi quadrati), con un'equazione di stima per ogni riga del modello. In questo caso l'equazione di stima non deve includere la costante (intercetta). Queste stime, ottenute con regressioni separate per ogni riga del modello, corrispondono a quelle che si otterrebbero minimizzando la somma dei quadrati delle componenti stocastiche di ogni riga, su tutte le righe del modello e su tutti i casi.

Se noi chiamiamo con E_{ik} la componente stocastica per il caso i e la riga k del modello, allora:

$$\begin{aligned} E_{ik} &= Y_{ik} - (g_{k1} X_{i1} + g_{k2} X_{i2} + \dots + g_{ki} X_{ii} + \dots + g_{kj} X_{ij}) = \\ &= d_{ik1} X_{i1} + d_{ik2} X_{i2} + \dots + d_{iki} X_{ii} + \dots + d_{ikj} X_{ij} = \sum_j d_{ikj} X_{ij} \end{aligned}$$

e la quantità che le stime minimizzano è data da:

$$\sum_i \sum_k E_{ik}^2$$

In altre parole, stimare ogni riga del modello separatamente con i metodi standard di regressione equivale a minimizzare la somma di tutte le componenti stocastiche per l'intero modello.

Le stime dei g_{ki} ottenute in questo modo, i \hat{g}_{ki} , fanno somma 1:

$$\sum_k \hat{g}_{ki} = 1$$

il che significa che le stime soddisfano almeno ad uno dei vincoli logici imposti dal modello ai coefficienti.

Tuttavia non necessariamente si verifica che le stime \hat{g}_{ki} siano comprese tra 0 ed 1, che è un'altra delle condizioni imposte dall'interpretazione che noi diamo ad esse in termini di frazione di popolazione.

Se l'assunto che abbiamo fatto, che in ogni unità territoriale ogni coefficiente differisca dal valore medio (i g_{ki}) di una piccola quantità stocastica (d_{ki}) è falso, allora esistono delle differenze sistematiche fra i coefficienti b_{ikj} delle diverse unità territoriali. Dal momento che queste differenze sono sistematiche, possono essere trattate mediante una loro inclusione nelle ipotesi relative ai coefficienti b :

$$b_{ikj} = g_{kj} + F_k(Z_i) + d_{ikj}$$

dove $F_k(Z_i)$ rappresenta una funzione F_k di un fattore sistematico Z che influenza il valore del coefficiente b nell'unità territoriale. Nel caso in cui tale funzione F sia una semplice funzione lineare, i coefficienti b_{ikj} sono dati da:

$$b_{ikj} = g_{kj} + c_{ki} Z_i + d_{ikj}$$

dove d_{ikj} è la piccola componente stocastica di prima, e c_{ki} è il coefficiente che lega il valore di Z nell'unità territoriale ai coefficienti b del modello. Poiché in ogni unità territoriale abbiamo che

$$\sum_k b_{ikj} = 1 = \sum_k (g_{kj} + c_{ki} Z_i + d_{ikj})$$

e poiché i valori Z_i variano da un'unità territoriale all'altra, questa condizione può essere soddisfatta solo se

$$\sum_k g_{kj} = 1$$

$$\sum_k c_{ki} Z_i = 0 \text{ per qualsiasi valore di } Z, \text{ per cui } \sum_k c_{ki} = 0$$

$$\sum_k d_{ikj} = 0$$

Possiamo quindi scrivere una riga del modello per il caso i:

$$Y_{ik} = \sum_j b_{ikj} X_{ij} = \sum_j g_{kj} X_{ij} + \sum_j c_{kj} Z_j X_{ij} + \sum_j d_{ikj} X_{ij}$$

I metodi standard di regressione di questo modello forniscono stime dei coefficienti g_{kj} (stima = \hat{g}_{kj}) e c_{kj} (stima = \hat{c}_{kj}), che hanno le proprietà:

$$\sum_j \hat{g}_{kj} = 1$$

$$\sum_j \hat{c}_{kj} = 0$$

È tuttavia possibile che non ci siano variabili Z disponibili, oppure che quelle disponibili non siano sufficienti o adatte per eliminare i coefficienti inaccettabili. In tal caso si possono usare metodi di regressione vincolata per ottenere dei coefficienti b accettabili. È chiaro che il ricorso a tali regressioni vincolate, oppure ad altri metodi di correzione dei risultati, è conseguenza di mancanza di teoria o di informazioni (non si sa quali sono le influenze sistematiche sui coefficienti b nelle varie unità territoriali, oppure esse sono note ma si è incapaci di misurarle).

È possibile misurare quanto le stime ottenute con le regressioni vincolate o con i metodi di correzione differiscono dalle stime ottenute coi metodi standard di regressione (criterio dei minimi quadrati). Il principio su cui poggia tale misura è il seguente. Supponiamo di aver ottenuto le stime dei coefficienti b del modello, e che alcune di queste siano negative. La maggior parte degli algoritmi di correzione, prima eguaglia a zero i coefficienti negativi, e poi corregge gli altri coefficienti per ottenere che siano soddisfatte le condizioni imposte ai coefficienti dal modello. Analogamente procedono in generale le tecniche di regressione vincolata, specialmente quando i vincoli sono dati, come nel nostro caso, dal fatto che i coefficienti devono fare somma 1 in ogni colonna e che quelli negativi vanno eguagliati a zero, con conseguente nuova stima dell'intero modello.

Siano \hat{b}_{ki} i coefficienti ottenuti con le tecniche standard di regressione, e b_{ki}^* i coefficienti derivanti invece dalle procedure di correzione o dalle regressioni vincolate. Ora i \hat{b}_{ki} che sono positivi, saranno in un certo senso «troppo grandi», a causa della presenza dei coefficienti negativi (si ricordi che i \hat{b}_{ki} fanno somma 1 in ogni colonna). Quindi una misura della differenza esistente fra i coefficienti prima e dopo l'applicazione dell'algoritmo di correzione o delle regressioni vincolate, è data dalla differenza fra i \hat{b}_{ki} positivi ed i corrispondenti b_{ki}^* corretti, moltiplicati per la frazione di popolazione alla quale essi si riferiscono (X_i).

In termini più precisi: sia q_{ki} uguale ad 1 quando $\hat{b}_{ki} > 0$, e 0 quando $\hat{b}_{ki} \leq 0$. Allora:

$$\sum \hat{b}_{ki} X_i \leq \sum q_{ki} \hat{b}_{ki} X_i$$

poiché il termine q_{ki} pone eguali a zero tutti i contributi ne-

gativi. Possiamo allora calcolare il coefficiente VR (*Valore Redistribuito*), che per un certo insieme di stime è dato da:

$$VR = \sum q_{kj} \hat{b}_{kj} X'_j - \sum b_{ki}^* X'_i$$

dove la sommatoria viene effettuata su tutti gli elementi dell'insieme considerato. Nella formula viene usato il valore X'_j invece di X_j , poiché stiamo trattando con stime che si riferiscono a più unità territoriali, e quindi il valore X'_j costituisce una sorta di media delle frazioni di popolazione che in ogni unità territoriale appartengono alla categoria j della variabile X . Analoga osservazione può essere fatta per il valore Y'_k che utilizzeremo più avanti.

Non è necessario, per il calcolo del coefficiente VR, tenere conto di tutti i coefficienti inaccettabili: è sufficiente considerare solo quelli negativi. Infatti i b_{kj} che sono maggiori di 1, sono tali solo perché ci sono dei coefficienti negativi, ed in questo senso non differiscono dagli altri coefficienti positivi, che pure sono «troppo grandi» a causa della presenza dei negativi.

In linea generale i coefficienti VR dipendono dal particolare tipo di procedura utilizzata per avere stime accettabili dei coefficienti b . Tuttavia i valori di alcuni dei possibili VR (si tratta peraltro dei casi in pratica più utilizzati) sono indipendenti dalla *routine* di correzione dei b . Infatti, dal momento che i b_{ki}^* soddisfano alle limitazioni del modello, abbiamo che:

$$\sum_i b_{ki}^* X'_i = X'_k$$

$$\sum_j b_{kj}^* X'_j = Y'_k$$

$$\sum_i \sum_k b_{ki}^* X'_i = \sum_k \sum_j b_{kj}^* X'_j = 1$$

Di conseguenza:

$$VR_{\text{riga } k} = \sum_j q_{kj} \hat{b}_{kj} X'_j - Y'_k$$

$$VR_{\text{colonna } j} = \sum_k q_{kj} \hat{b}_{kj} X'_j - X'_j$$

$$VR_{\text{totale}} = \sum_j \sum_k q_{kj} \hat{b}_{kj} X'_j - 1 = \sum_j VR_{\text{colonna } j} = \sum_k VR_{\text{riga } k}$$

Si può vedere che questi VR in nessun modo dipendono dai valori dei singoli coefficienti b_{ki}^* . I coefficienti VR relativi alla correzione totale, così come quelli relativi alla correzione per riga o colonna, sono indipendenti dalle modalità scelte per la correzione e dipendono solo dalle stime ottenute tramite le regressioni condotte sui dati. Ciò è importante poiché significa che i coefficienti VR forniscono una misura diretta del grado con il quale il modello si adatta ai dati.

Oltre a quelli citati (per riga, per colonna e sul totale), si possono calcolare altri coefficienti VR, di minor interesse, però. Ce n'è uno tuttavia sul quale torneremo. Nel caso in cui la variabile Y è il voto di una elezione, e la variabile X il voto nella elezione precedente, gli elettori che votano due volte per lo stesso partito si collocano nelle celle $\hat{b}_{kj} X'_j$ dove $k = j$ (cioè sulla diagonale), per cui:

$$VR_{\text{diagonale}} = \sum_k q_{kk} \hat{b}_{kk} X'_k - \sum_k b_{ki}^* X'_i \quad \text{per } k = j$$

Si può notare che in questo caso VR dipende dalla particolare procedura di correzione dei b utilizzata.

I coefficienti VR sono interpretabili in maniera abbastanza naturale. Dal momento che i valori b_i, X_i rappresentano, come abbiamo visto, delle frazioni di popolazione, anche i coefficienti VR possono essere interpretati come «frazioni di popolazione». Essi costituiscono quella parte di popolazione che è stata «rimossa» dal totale per rendere le stime dei coefficienti b accettabili nei termini dei vincoli logici ad essi imposti dal modello. Oltre a ciò i coefficienti VR danno un'idea, anche se piuttosto rozza, di quanto mediamente le stime positive dei coefficienti b devono «contrarsi» perché si possa arrivare ad un insieme di stime che soddisfi ai vincoli logici del modello. Valori del coefficiente VR_{totale} superiori a 0.15 (quando la popolazione è scritta in termini di frazioni, o sopra 15 quando la popolazione è scritta in termini di percentuali) ci sembrano, in generale, inaccettabili; ma ogni valore di VR_{totale} superiore a 0.10 deve essere guardato con un certo sospetto.

Al di là del problema dei coefficienti inaccettabili, ci possiamo chiedere quanto il modello effettivamente è in grado di prevedere il voto e la sua variabilità. L'ovvia misura per questo problema è data dal quadrato del coefficiente di correlazione di ogni riga del modello, R^2 , che ci dà, per ogni regressione, la frazione di varianza della variabile dipendente spiegata dalle variabili indipendenti. Quando nel testo parliamo di coefficienti di correlazione, intendiamo questa misura. Una misura equivalente per l'intero modello potrebbe essere data dalla somma ponderata e standardizzata dei K coefficienti di correlazione R^2 , dove le varianze delle variabili indipendenti agirebbero come fattori di ponderazione. Se noi chiamiamo R^2 (R^2 totale) questo coefficiente ed usiamo la notazione $Var(\cdot)$ per la varianza di una variabile, abbiamo:

$$R^2 = \frac{\sum R_i^2 Var(Y_k)}{\sum Var(Y_k)}$$

Il coefficiente R^2 è una media ponderata dei coefficienti R_i^2 ; il suo valore sarà influenzato di più dalle variabili dipendenti (partiti) che hanno maggiore varianza (nel loro voto), rispetto a quelle con varianza inferiore. Ciò significa in pratica che gli R_i^2 dei maggiori partiti tendono ad avere un impatto maggiore dei piccoli partiti sul valore finale di R^2 .

Tuttavia un altro indice può assumere un certo interesse nello stabilire quanto il modello «spiega» del voto. Questo poiché non solo la varianza del voto, ma anche il suo effettivo livello, è degno di interesse. Quindi, invece di chiedere quale frazione della varianza è spiegata dalla regressione, ci si può chiedere quale parte della somma dei quadrati della variabile dipendente è spiegata dalla regressione.

Quando l'equazione di regressione contiene una costante:

$$Y = c + \sum b_i X_i + E$$

l'informazione è di scarso interesse poiché la costante può essere stimata attraverso:

$$\hat{c} = Me(Y) - \sum Me(b_i X_i)$$

dove Me sta per l'operatore che calcola la media di una variabile. Ciò che viene qui mostrato è che la media (osservata) della variabile dipendente coincide sempre con la media predetta, ed anche che la media della componente stocastica, $Me(E)$, è 0. Di conseguenza, se chiamiamo con r^2 la frazione della somma totale dei quadrati della variabile dipendente spiegata dalla regressione, troviamo:

$$r^2 = (\Sigma Y^2 - \Sigma E^2) / \Sigma Y^2$$

ma pure abbiamo che

$$R^2 = (\text{Var}(Y) - \text{Var}(E)) / \text{Var}(Y)$$

e che

$$\Sigma Y^2 = N(\text{Var}(Y) + \text{Me}(Y)^2)$$

dove N è il numero di casi sui quali lavoriamo. Inoltre

$$\Sigma E^2 = N(\text{Var}(E) + \text{Me}(E)^2) = N \text{Var}(E)$$

dal momento che $\text{Me}(E) = 0$ come abbiamo già visto. E dalla formula di R^2 ricaviamo:

$$R^2 N \text{Var}(Y) = N(\text{Var}(Y) - \text{Var}(E)),$$

quindi

$$\begin{aligned} r^2 &= (\Sigma Y^2 - \Sigma E^2) / \Sigma Y^2 = \\ &= (N \text{Var}(Y) - N \text{Var}(E) + N \text{Me}(Y)^2) / (N(\text{Var}(Y) + \text{Me}(Y)^2)) \\ r^2 &= (R^2 \text{Var}(Y) + \text{Me}(Y)^2) / (\text{Var}(Y) + \text{Me}(Y)^2) \end{aligned}$$

che mostra che c'è una semplice relazione fra r^2 e R^2 , espressa in termini di varianza e di media della variabile Y, quando la regressione include una costante. Ma questa relazione non è più necessariamente vera quando l'equazione non include una costante, e questa è la ragione per la quale, in questo caso, molti programmi di regressione forniscono il coefficiente r^2 invece del R^2 ¹.

Se si controlla la dimostrazione data sopra, si può vedere che la parte cruciale del ragionamento è quella che presuppone che la media della componente stocastica sia zero, cioè $\text{Me}(E) = 0$. Infatti, tutte le volte che l'equazione include una costante, si può sempre garantire che la somma delle componenti stocastiche faccia 0, poiché la costante può essere scelta in modo da ottenere questo risultato. Senza costante ciò non è sempre possibile. Ma nello speciale caso considerato quest'ultima affermazione non tiene. Dal momento che $\Sigma X_i = 1$, noi possiamo sempre riscrivere le equazioni in modo tale che esse contengano una costante, mantenendo inalterata la componente stocastica:

poiché

$$X_1 = 1 - \Sigma_j X_j \quad j = 2, \dots, J$$

allora

$$\begin{aligned} Y_k &= b_{k1} X_1 + \sum_j b_{kj} X_j + E_k & j &= 2, \dots, J \\ &= b_{k1} - b_{k1} \sum_j X_j + \sum_j b_{kj} X_j + E_k & j &= 2, \dots, J \\ &= b_{k1} + \sum_j (b_{kj} - b_{k1}) X_j + E_k & j &= 2, \dots, J \end{aligned}$$

¹ Nella notazione utilizzata dal programma SPSS, quando si utilizza una regressione forzata attraverso l'origine (cioè senza costante), quello che noi abbiamo chiamato r^2 corrisponde a R SQUARE riportato senza altre precisazioni, mentre il nostro R^2 corrisponde al R SQUARE riportato sotto la dicitura «total sum of squares adjusted for the mean of dependent variable».

Da ciò segue che $Me(E_k) = 0$, e quindi la relazione prima fornita fra R^2 e r^2 è verificata anche per il modello qui considerato. Quando $Me(E) = 0$, r^2 assume un minor interesse, poiché conoscendo R^2 , che è il normale coefficiente di correlazione, la media e la varianza della variabile dipendente, r^2 può essere immediatamente calcolato.

Possiamo osservare dalla formula che $r^2 \geq R^2$, con l'eguaglianza solo nel caso in cui $Me(Y) = 0$, il che è escluso dai dati per il modello che stiamo qui considerando. Quando la media della variabile dipendente è abbastanza grande rispetto alla varianza, il coefficiente r^2 sarà sempre prossimo a 1. Si può costruire una misura complessiva della bontà dell'adattamento dell'intero modello, utilizzando i coefficienti r_k^2 , mediante una somma ponderata e standardizzata che utilizza la somma dei quadrati delle osservazioni come fattore di ponderazione. La formula per questo r^2 complessivo e ponderato è la seguente:

$$r^2 = \frac{\sum_k (r_k^2 \sum_i Y_{ik})}{\sum_k \sum_i \bar{Y}_{ik}}$$

che è parallela alla formula data per il coefficiente R^2 .

In generale, per il modello che stiamo considerando, il coefficiente r^2 è di interesse minore rispetto al coefficiente R^2 . Ma c'è un'eccezione a questa regola. Consideriamo la situazione in cui, in prima istanza, la stima dei coefficienti b_k tramite i minimi quadrati ordinari ha portato a coefficienti negativi inaccettabili, e nella quale non siano disponibili variabili addizionali in grado di condurre a stime accettabili. In questo caso, come s'è detto, quando i valori negativi non sono troppo alti, oppure, nei termini del coefficiente VR qui proposto, quando il coefficiente VR non è troppo alto, si può ricorrere a regressioni vincolate oppure ad una *routine* di correzione. Nella maggior parte degli approcci basati su regressioni vincolate, i coefficienti negativi b_{k1} vengono eguagliati a zero, o, detto in altri termini, la variabile X_{k1} viene esclusa dalle regressioni di stima di Y_{k1} . Ma in tal caso non c'è più alcuna garanzia che $Me(E_k) = 0$, e quindi la relazione fra r^2 e

R^2 , prima data, non è più valida. Inoltre in questo caso non è neppure più valido che $\sum E_k^2 = N(\text{Var}(E_k) + \text{Me}(E_k)) = N\text{Var}(E_k)$, per cui l'uso della formula di R^2 , che si basa sulle varianze, può condurre ad un'ingannevole impressione sulla bontà dell'adattamento dell'intero modello. Il coefficiente r^2 , che è formulato in termini di somme, non risente di questo problema, e dovrebbe essere usato in questi casi, anche se siamo meno familiari con esso, e quindi la sua interpretazione può portare qualche problema.

Quando le stime dei coefficienti b del modello sono ottenute coi metodi dei minimi quadrati, sia minimi quadrati ordinari, che minimi quadrati vincolati o minimi quadrati vincolati in regressioni apparentemente indipendenti (*seemingly unrelated regressions* [cfr. Zellner 1962]), i residui hanno la proprietà che:

$$\sum_i X_{ij} E_{ik} = 0$$

che permette la suddivisione della somma dei quadrati totale della variabile indipendente in due parti indipendenti l'una dall'altra.

Se

$$Y_k = \sum_i b_{ki} X_i + E_k$$

allora

$$Y_k - E_k = \sum_i b_{ki} X_i = \hat{Y}$$

quindi

$$\begin{aligned} Y_k^2 &= \sum_i (\hat{Y}_i + E_{ik})^2 = \sum_i \hat{Y}_i^2 + \sum_i E_{ik}^2 + 2 \sum_i \sum_j (b_{ki} X_{ij}) E_{jk} = \\ &= \sum_i \hat{Y}_i^2 + \sum_i E_{ik}^2 \end{aligned}$$

che ci fornisce una definizione di r_k^2 non ambigua. Ma nelle formule presentate in questo testo viene utilizzata un'altra *routine* di correzione, che ha come conseguenza il fatto che i coefficienti b che sono ottenuti, non sono più stime dei minimi quadrati. Questo algoritmo presenta il vantaggio che le medie delle variabili dipendenti predette sulla base del modello sono uguali a quelle osservate, cioè

$$\text{Me}(Y) = \text{Me}(\hat{Y}) \quad \text{e} \quad \text{Me}(E_k) = 0$$

ma non è più vero che

$$\sum_i X_{ij} E_{ik} = 0$$

Di conseguenza la somma totale dei quadrati delle osservazioni, $\sum_i Y_{ik}^2$ non può essere più divisa in due parti, una dovuta solo alla regressione, e l'altra solo alla componente stocastica; rimane un elemento che è dovuto alla combinazione di entrambe (cioè $\sum_i X_{ij} E_{jk}$) che non può essere assegnato con certezza alla variabile dipendente od alla componente stocastica. Di conseguenza gli r^2 e gli R^2 non possono essere definiti in maniera univoca dopo l'applicazione dell'algoritmo di correzione. Per questa ragione i coefficienti R^2 e (talvolta r^2) che presentiamo, e con loro i coefficienti complessivi R^2 e r^2 , si riferiscono alle stime ottenute prima dell'applicazione della *routine* di correzione, e non dopo.

Si può notare che quando non ci sono coefficienti inaccettabili, e quindi non ci sono correzioni, i coefficienti R^2 e r^2 possono essere utilizzati per ottenere un'idea diretta della bontà dell'adattamento. Se invece è necessario eliminare dei coefficienti inaccettabili, e l'operazione viene effettuata col solo aiuto di regressioni vincolate, soltanto i coefficienti r^2 possono essere utilizzati per stabilire la bontà dell'adattamento dopo le correzioni; mentre se la correzione avviene tramite la *routine* proposta nel testo (cfr. capitolo settimo), né R^2 né r^2 possono essere utilizzati dopo la correzione. In ogni caso la correzione dei coefficienti inaccettabili sarà fatta solo se il loro numero e la loro dimensione saranno di modesta entità;

ed in tal caso le variazioni della componente saranno piccole ed i coefficienti R^2 ed r^2 ottenuti prima della correzione costituiranno un buon indicatore dell'adattamento globale del modello, anche dopo correzione (per un ulteriore approfondimento sulle varie procedure di correzione dei coefficienti inaccettabili si veda il capitolo settimo).

Per concludere. In questo paragrafo abbiamo presentato un modello di analisi dei dati aggregati nel quale le variabili dipendenti ed indipendenti rappresentano delle frazioni di popolazione; il che permette l'interpretazione dei coefficienti b stimati attraverso i metodi di regressione, in termini di frazioni di popolazione. Di conseguenza i coefficienti b non possono essere negativi; e la somma dei coefficienti che si riferiscono alla stessa variabile indipendente (nelle diverse equazioni di regressione) deve fare somma 1. Le stime ordinarie dei minimi quadrati posseggono la seconda proprietà, ma non necessariamente la prima. Nel caso si presentino delle stime inaccettabili, occorre cercare le variabili aggiuntive che agiscono in maniera sistematica sui coefficienti b . Se non vengono trovate, occorre ricorrere a procedure di correzione per eliminare i coefficienti inaccettabili; tale ricorso tuttavia è giustificato solo se le correzioni da fare sono piccole. A questo proposito abbiamo suggerito una misura, il coefficiente VR. Infine, l'adattamento totale del modello può essere misurato mediante i quadrati degli usuali coefficienti di correlazione e mediante un coefficiente che rappresenta la frazione della somma dei quadrati della variabile dipendente spiegata dalle variabili indipendenti.

ULTIMI WORKING PAPERS

- 64 "L'attività in agricoltura e il censimento demografico del 1981", maggio 1985
- 65 "Stima della struttura dei consumi familiari commercializzati a scala sub-regionale", marzo 1985
- 66 "Simulazione dell'impatto di scenari socio-economici e di politiche di trasporto sul sistema urbano di Torino", maggio 1985
- 67 "Elaborazione dei dati censuari sulle attività commerciali a base comunale, con aggregazione a livello comprensoriale", maggio 1985
- 68 "Lo sviluppo di una procedura computerizzata interattiva per la pianificazione sanitaria regionale", giugno 1985
- 69 "L'evoluzione delle gerarchie territoriali in Piemonte", giugno 1985
- 70 "An integrated model for the dynamic analysis of location-transport interrelation", luglio 1985
- 71 "L'Agricoltura piemontese nel 1984 attraverso i dati dell'Osservatorio Contabile Regionale (O.C.R.), aprile 1986
- 72 "Livello e qualità della vita in Piemonte", aprile 1986
- 73 "Valutazione delle quote di mercato e dei livelli di modernizzazione del sistema distributivo alimentare per aree subregionali, dicembre 1986
- 74 "Se io fossi il Sindaco... Le preferenze fiscali prese sul serio. Rapporto di ricerca sulle preferenze fiscali a Torino, dicembre 1986
- 75 "Utilizzo della domanda pubblica regionale ai fini della promozione tecnologica e produttiva di alcuni settori in Piemonte", aprile 1987
- 76 "Industria e innovazione - L'area dell'automazione industriale", luglio 1987
- 77 "Elaborati conoscitivi e metodologici dell'Osservatorio demografico territoriale", luglio 1987
- 78 "Studi sulla marginalità in agricoltura in un'area del Piemonte. L'agricoltura del comprensorio di Mondovì attraverso i censimenti e le analisi aziendali", ottobre 1987
- 79 "L'occupazione nella pubblica amministrazione negli anni '80: tendenze e prospettive", novembre 1987
- 80 "Il part-time nella Pubblica Amministrazione: problemi e prospettive", novembre 1987
- 81 "Revealed preferences for local public goods: the Turin experiment", dicembre 1987

- 82 "Il problema dei flussi scolastici: un modello di analisi", dicembre 1987
- 83 "L'agricoltura a tempo parziale in Piemonte: un'analisi dei dati del III Censimento generale dell'agricoltura", marzo 1988
- 84 "Archivio degli indicatori sociali. Un approccio costruttivista all'organizzazione dei dati", aprile 1988

L'IRES è stato costituito nel 1958 dalla Provincia e dal Comune di Torino, con la partecipazione di altri enti pubblici e privati. Con la successiva adesione delle altre Province piemontesi, l'Istituto ha assunto carattere regionale.

Nel 1974 l'IRES è diventato ente strumentale della Regione Piemonte ed è stato dotato di personalità giuridica di diritto pubblico.

L'attività dell'IRES è attualmente disciplinata dalla legge regionale 18 febbraio 1985, n. 12.

L'IRES, struttura primaria di ricerca della Regione Piemonte, sviluppa la propria attività in raccordo con le esigenze della azione programmatoria ed operativa della Regione stessa, degli Enti locali e degli Enti pubblici.

Costituiscono oggetto dell'attività dell'Istituto:

- la redazione della relazione annuale sull'andamento socio-economico e territoriale della Regione;
- la conduzione di una permanente attività di osservazione, documentazione ed analisi sulle principali grandezze socio-economiche e territoriali del sistema regionale;
- lo svolgimento di periodiche rassegne congiunturali sull'economia regionale;
- lo svolgimento delle ricerche connesse alla redazione ed alla attuazione del piano regionale di sviluppo;
- lo svolgimento di ricerche di settore per conto della Regione e altri enti.

ires

ISTITUTO RICERCHE ECONOMICO - SOCIALI DEL PIEMONTE
VIA BOGINO 21 10123 TORINO